SUBSTITUTE SEQUENCE LISTING

<110> Dumas Milne Edwards, Jean-Baptiste
      Duclert, Aymeric
      Bougueleret, Lydie

<120> EXTENDED CDNAS FOR SECRETED PROTEINS

<130> 31.US3.CIP

<140> 09/663,600
<141> 2000-09-15

<150> 09/191,997
<151> 1998-11-13

<150> 60/066,677
<151> 1997-11-13

<150> 60/069,957
<151> 1997-12-17

<150> 60/074,121
<151> 1998-02-09

<150> 60/081,563
<151> 1998-04-13

<150> 60/096,116
<151> 1998-08-10

<150> 60/099,273
<151> 1998-09-04

<160> 229

<170> Patent.pm

<210> 1
<211> 47
<212> RNA
<213> Artificial Sequence
<220>
<223> in vitro transcription product
<220>
<221> modified_base
<222> 1
<223> m7g
<400> 1
ggcauccuac ucccauccaa uuccacccua acuccuccca ucuccac   47

<210> 2
<211> 46
<212> RNA
<213> Artificial Sequence
<220>
<223> in vitro transcription product

1

<400> 2
gcauccuacu cccauccaau uccacccuaa cuccucccau cuccac   46

<210> 3
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Derivatized oligonucleotide for linking to mRNA
<400> 3
atcaagaatt cgcacgagac catta   25

<210> 4
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Probe complementary to oligonucleotide of SEQ ID NO:3
<400> 4
taatggtctc gtgcgaattc ttgat   25

<210> 5
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Alpha globin gene primer GLO-S
<400> 5
ccgacaagac caacgtcaag gccgc   25

<210> 6
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Alpha globin gene primer GLO-As
<400> 6
tcaccagcag gcagtggctt aggag   25

<210> 7
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Dehydrogenase gene primer 3 DH-S
<400> 7
agtgattcct gctactttgg atggc   25

<210> 8
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Dehydrogenase gene primer 3 DH-As
<400> 8
gcttggtctt gttctggagt ttaga   25

<210> 9

2

```
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> PP15 gene primer PP15-S
<400> 9
tccagaatgg gagacaagcc aattt                                              25


<210> 10
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> PP15 gene primer PP15-As
<400> 10
agggaggagg aaacagcgtg agtcc                                              25


<210> 11
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Elongation Factor A4 gene primer EFA1-S
<400> 11
atgggaaagg aaaagactca tatca                                              25


<210> 12
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Elongation Factor A4 gene primer EF1A-As
<400> 12
agcagcaaca atcaggacag cacag                                              25


<210> 13
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Primer derived from oligonucleotide of SEQ ID NO:3
<400> 13
atcaagaatt cgcacgagac catta                                              25


<210> 14
<211> 67
<212> DNA
<213> Artificial Sequence
<220>
<223> PolydT primer
<220>
<221> misc_feature
<222> 67
<223> n=a, g, c or t
<400> 14
atcgttgaga ctcgtaccag cagagtcacg agagagacta cacggtactg gttttttttt      60
tttttvn                                                                 67
```

3

<210> 15
<211> 29
<212> DNA
<213> Artificial Sequence
<220>
<223> Nested 3' primer
<400> 15
ccagcagagt cacgagagag actacacgg   29

<210> 16
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> Nested 3' primer
<400> 16
cacgagagag actacacggt actgg   25

<210> 17
<211> 526
<212> DNA
<213> Homo Sapiens
<220>
<221> misc_feature
<222> complement(261..376)
<223> blastn
<220>
<221> misc_feature
<222> complement(380..486)
<223> blastn
<220>
<221> misc_feature
<222> complement(110..145)
<223> blastn
<220>
<221> misc_feature
<222> complement(196..229)
<223> blastn
<220>
<221> sig_peptide
<222> 90..140
<223> Von Heijne matrix
<220>
<221> misc_feature
<222> 290
<223> n=a, g, c or t
<400> 17
aatatrarac agctacaata ttccagggcc artcacttgc catttctcat aacagcgtca   60
gagagaaaga actgactgar acgtttgag atg aag aaa gtt ctc ctc ctg atc   113
                                 Met Lys Lys Val Leu Leu Leu Ile
                                         -15   -10
aca gcc atc ttg gca gtg gct gtw ggt ttc cca gtc tct caa gac cag   161
Thr Ala Ile Leu Ala Val Ala Val Gly Phe Pro Val Ser Gln Asp Gln
             -5   1   5
gaa cga gaa aaa aga agt atc agt gac agc gat gaa tta gct tca ggr   209
Glu Arg Glu Lys Arg Ser Ile Ser Asp Ser Asp Glu Leu Ala Ser Gly
         10   15   20
wtt ttt gtg ttc cct tac cca tat cca ttt cgc cca ctt cca cca att   257

4

```
Xaa Phe Val Phe Pro Tyr Pro Tyr Pro Phe Arg Pro Leu Pro Pro Ile
    25                      30                  35
cca ttt cca aga ttt cca tgg ttt aga cgt aan ttt cct att cca ata          305
Pro Phe Pro Arg Phe Pro Trp Phe Arg Arg Xaa Phe Pro Ile Pro Ile
40                      45                  50                  55
cct gaa tct gcc cct aca act ccc ctt cct agc gaa aag taaacaaraa          354
Pro Glu Ser Ala Pro Thr Thr Pro Leu Pro Ser Glu Lys
                    60                  65
ggaaaagtca crataaacct ggtcacctga aattgaaatt gagccacttc cttgaaraat          414
caaaattcct gttaataaaa raaaaacaaa tgtaattgaa atagcacaca gcattctcta          474
gtcaatatct ttagtgatct tctttaataa acatgaaagc aaaaaaaaaa aa          526
```

```
<210> 18
<211> 17
<212> PRT
<213> Homo Sapiens
<220>
<221> SIGNAL
<222> 1..17
<223> Von Heijne matrix
      score 8.2
      seq LLLITAILAVAVG/FP
<400> 18
Met Lys Lys Val Leu Leu Leu Ile Thr Ala Ile Leu Ala Val Ala Val
1               5                   10                  15
Gly
```

```
<210> 19
<211> 822
<212> DNA
<213> Homo Sapiens
<220>
<221> misc_feature
<222> 260..464
<223> blastn
<220>
<221> misc_feature
<222> 118..184
<223> blastn
<220>
<221> misc_feature
<222> 56..113
<223> blastn
<220>
<221> misc_feature
<222> 454..485
<223> blastn
<220>
<221> misc_feature
<222> 118..545
<223> blastn
<220>
<221> misc_feature
<222> 65..369
<223> blastn
<220>
<221> misc_feature
<222> 61..399
```

5

<223> blastn
<220>
<221> misc_feature
<222> 408..458
<223> blastn
<220>
<221> misc_feature
<222> 60..399
<223> blastn
<220>
<221> misc_feature
<222> 393..432
<223> blastn
<220>
<221> sig_peptide
<222> 346..408
<223> Von Heijne matrix
<220>
<221> misc_feature
<222> 115
<223> n=a, g, c or t
<400> 19

```
actcctttta gcatagggc ttcggcgcca gcggccagcg ctagtcggtc tggtaagtgc      60
ctgatgccga gttccgtctc tcgcgtcttt tcctggtccc aggcaaagcg gasgnagatc     120
ctcaaacggc ctagtgcttc gcgcttccgg agaaaatcag cggtctaatt aattcctctg     180
gtttgttgaa gcagttacca agaatcttca accctttccc acaaaagcta attgagtaca     240
cgttcctgtt gagtacacgt tcctgttgat ttacaaaagg tgcaggtatg agcaggtctg     300
aagactaaca ttttgtgaag ttgtaaaaca gaaaacctgt tagaa atg tgg tgg ttt     357
                                                Met Trp Trp Phe
                                                -20

cag caa ggc ctc agt ttc ctt cct tca gcc ctt gta att tgg aca tct      405
Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr Ser
        -15               -10               -5

gct gct ttc ata ttt tca tac att act gca gta aca ctc cac cat ata      453
Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His Ile
 1             5                   10                  15

gac ccg gct tta cct tat atc agt gac act ggt aca gta gct cca raa      501
Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro Xaa
             20                  25                  30

aaa tgc tta ttt ggg gca atg cta aat att gcg gca gtt tta tgt caa      549
Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys Gln
         35                  40                  45

aaa tagaaatcag gaarataatt caacttaaag aakttcattt catgaccaaa           602
Lys
ctcttcaraa acatgtcttt acaagcatat ctcttgtatt gctttctaca ctgttgaatt    662
gtctggcaat atttctgcag tggaaaattt gatttarmta gttcttgact gataaatatg    722
gtaaggtggg ctttttcccc tgtgtaattg ctactatgt cttactgagc caagttgtaw    782
tttgaaataa aatgatatga gagtgacaca aaaaaaaaaa                          822
```

<210> 20
<211> 21
<212> PRT
<213> Homo Sapiens
<220>
<221> SIGNAL
<222> 1..21
<223> Von Heijne matrix
        score 5.5

6

```
          seq SFLPSALVIWTSA/AF
<400> 20
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
1               5                   10                  15
Ile Trp Thr Ser Ala
            20


<210> 21
<211> 405
<212> DNA
<213> Homo Sapiens
<220>
<221> misc_feature
<222> complement(103..398)
<223> blastn
<220>
<221> sig_peptide
<222> 185..295
<223> Von Heijne matrix
<400> 21
atcaccttct tctccatcct tstctgggcc agtccccarc ccagtccctc tcctgacctg        60
cccagcccaa gtcagccttc agcacgcgct tttctgcaca cagatattcc aggcctacct       120
ggcattccag gacctccgma atgatgctcc agtcccttac aagcgcttcc tggatgaggg       180
tggc atg gtg ctg acc acc ctc ccc ttg ccc tct gcc aac agc cct gtg       229
     Met Val Leu Thr Thr Leu Pro Leu Pro Ser Ala Asn Ser Pro Val
              -35                 -30                 -25
aac atg ccc acc act ggc ccc aac agc ctg agt tat gct agc tct gcc       277
Asn Met Pro Thr Thr Gly Pro Asn Ser Leu Ser Tyr Ala Ser Ser Ala
        -20                 -15                 -10
ctg tcc ccc tgt ctg acc gct cca aak tcc ccc cgg ctt gct atg atg       325
Leu Ser Pro Cys Leu Thr Ala Pro Xaa Ser Pro Arg Leu Ala Met Met
    -5                  1               5                   10
cct gac aac taaatatcct tatccaaatc aataaarwra raatcctccc               374
Pro Asp Asn
tccaraaggg tttctaaaaa caaaaaaaaa a                                     405


<210> 22
<211> 37
<212> PRT
<213> Homo Sapiens
<220>
<221> SIGNAL
<222> 1..37
<223> Von Heijne matrix
       score 5.9
       seq LSYASSALSPCLT/AP
<400> 22
Met Val Leu Thr Thr Leu Pro Leu Pro Ser Ala Asn Ser Pro Val Asn
1               5                   10                  15
Met Pro Thr Thr Gly Pro Asn Ser Leu Ser Tyr Ala Ser Ser Ala Leu
            20                  25                  30
Ser Pro Cys Leu Thr
            35


<210> 23
<211> 496
<212> DNA
<213> Homo Sapiens
```

<220>
<221> misc_feature
<222> 149..331
<223> blastn
<220>
<221> misc_feature
<222> 328..485
<223> blastn
<220>
<221> misc_feature
<222> complement(182..496)
<223> blastn
<220>
<221> sig_peptide
<222> 196..240
<223> Von Heijne matrix
<220>
<221> misc_feature
<222> 101
<223> n=a, g, c or t
<400> 23

```
aaaaaattgg tcccagtttt caccctgccg cagggctggc tggggagggc agcggtttag        60
attagccgtg gcctaggccg tttaacgggg tgacacgagc ntgcagggcc gagtccaagg       120
cccggagata ggaccaaccg tcaggaatgc gaggaatgtt tttcttcgga ctctatcgag       180
gcacacagac agacc atg ggg att ctg tct aca gtg aca gcc tta aca ttt        231
                  Met Gly Ile Leu Ser Thr Val Thr Ala Leu Thr Phe
                  -15                 -10                 -5
gcc ara gcc ctg gac ggc tgc aga aat ggc att gcc cac cct gca agt        279
Ala Xaa Ala Leu Asp Gly Cys Arg Asn Gly Ile Ala His Pro Ala Ser
        1                   5                   10
gag aag cac aga ctc gag aaa tgt agg gaa ctc gag asc asc cac tcg        327
Glu Lys His Arg Leu Glu Lys Cys Arg Glu Leu Glu Xaa Xaa His Ser
    15                  20                  25
gcc cca gga tca acc cas cac cga aga aaa aca acc aga aga aat tat        375
Ala Pro Gly Ser Thr Xaa His Arg Arg Lys Thr Thr Arg Arg Asn Tyr
30                  35                  40                  45
tct tca gcc tgaaatgaak ccgggatcaa atggttgctg atcaragccc              424
Ser Ser Ala
atatttaaat tggaaaagtc aaattgasca ttattaaata agcttgttt aatatgtctc       484
aaacaaaaaa aa                                                          496
```

<210> 24
<211> 15
<212> PRT
<213> Homo Sapiens
<220>
<221> SIGNAL
<222> 1..15
<223> Von Heijne matrix
       score 5.5
       seq ILSTVTALTFAXA/LD
<220>
<221> UNSURE
<222> 14
<223> Xaa = any one of the twenty amino acids
<400> 24

```
Met Gly Ile Leu Ser Thr Val Thr Ala Leu Thr Phe Ala Xaa Ala
1               5                   10                  15
```

8 137

<210> 25
<211> 623
<212> DNA
<213> Homo Sapiens
<220>
<221> sig_peptide
<222> 49..96
<223> Von Heijne matrix
<400> 25

```
aaagatccct gcagcccggc aggagagaag gctgagcctt ctggcgtc atg gag agg        57
                                                        Met Glu Arg
                                                            -15
ctc gtc cta acc ctg tgc acc ctc ccg ctg gct gtg gcg tct gct ggc        105
Leu Val Leu Thr Leu Cys Thr Leu Pro Leu Ala Val Ala Ser Ala Gly
            -10                 -5                      1
tgc gcc acg acg cca gct cgc aac ctg agc tgc tac cag tgc ttc aag        153
Cys Ala Thr Thr Pro Ala Arg Asn Leu Ser Cys Tyr Gln Cys Phe Lys
    5                   10                  15
gtc agc agc tgg acg gag tgc ccg ccc acc tgg tgc agc ccg ctg gac        201
Val Ser Ser Trp Thr Glu Cys Pro Pro Thr Trp Cys Ser Pro Leu Asp
20                  25                  30                      35
caa gtc tgc atc tcc aac gag gtg gtc gtc tct ttt aaa tgg agt gta        249
Gln Val Cys Ile Ser Asn Glu Val Val Val Ser Phe Lys Trp Ser Val
                40                  45                  50
cgc gtc ctg ctc agc aaa cgc tgt gct ccc aga tgt ccc aac gac aac        297
Arg Val Leu Leu Ser Lys Arg Cys Ala Pro Arg Cys Pro Asn Asp Asn
            55                  60                  65
atg aak ttc gaa tgg tcg ccg gcc ccc atg gtg caa ggc gtg atc acc        345
Met Xaa Phe Glu Trp Ser Pro Ala Pro Met Val Gln Gly Val Ile Thr
            70                  75                  80
agg cgc tgc tgt tcc tgg gct ctc tgc aac agg gca ctg acc cca cag        393
Arg Arg Cys Cys Ser Trp Ala Leu Cys Asn Arg Ala Leu Thr Pro Gln
            85                  90                  95
gag ggg cgc tgg gcc ctg cra ggg ggg ctc ctg ctc cag gac cct tcg        441
Glu Gly Arg Trp Ala Leu Xaa Gly Gly Leu Leu Leu Gln Asp Pro Ser
100                 105                 110                     115
agg ggc ara aaa acc tgg gtg cgg cca cag ctg ggg ctc cca ctc tgc        489
Arg Gly Xaa Lys Thr Trp Val Arg Pro Gln Leu Gly Leu Pro Leu Cys
                120                 125                 130
ctt ccc awt tcc aac ccc ctc tgc cca rgg gaa acc cag gaa gga            534
Leu Pro Xaa Ser Asn Pro Leu Cys Pro Xaa Glu Thr Gln Glu Gly
                135                 140                 145
taacactgtg ggtgccccca cctgtgcatt gggaccacra cttcaccctc ttggaracaa      594
taaactctca tgcccccaaa aaaaaaaaa                                        623
```

<210> 26
<211> 16
<212> PRT
<213> Homo Sapiens
<220>
<221> SIGNAL
<222> 1..16
<223> Von Heijne matrix
        score 10.1
        seq LVLTLCTLPLAVA/SA
<400> 26
Met Glu Arg Leu Val Leu Thr Leu Cys Thr Leu Pro Leu Ala Val Ala

```
<210> 27
<211> 848
<212> DNA
<213> Homo Sapiens
<220>
<221> sig_peptide
<222> 32..73
<223> Von Heijne matrix
<400> 27
aactttgcct tgtgttttcc accctgaaag a atg ttg tgg ctg ctc ttt ttt        52
                                    Met Leu Trp Leu Leu Phe Phe
                                                        -10
ctg gtg act gcc att cat gct gaa ctc tgt caa cca ggt gca gaa aat       100
Leu Val Thr Ala Ile His Ala Glu Leu Cys Gln Pro Gly Ala Glu Asn
             -5                1                   5
gct ttt aaa gtg aga ctt agt atc aga aca gct ctg gga gat aaa gca       148
Ala Phe Lys Val Arg Leu Ser Ile Arg Thr Ala Leu Gly Asp Lys Ala
10                  15                  20                  25
tat gcc tgg gat acc aat gaa gaa tac ctc ttc aaa gcg atg gta gct       196
Tyr Ala Trp Asp Thr Asn Glu Glu Tyr Leu Phe Lys Ala Met Val Ala
                 30                  35                  40
ttc tcc atg aga aaa gtt ccc aac aga gaa gca aca gaa att tcc cat       244
Phe Ser Met Arg Lys Val Pro Asn Arg Glu Ala Thr Glu Ile Ser His
             45                  50                  55
gtc cta ctt tgc aat gta acc cag agg gta tca ttc tgg ttt gtg gtt       292
Val Leu Leu Cys Asn Val Thr Gln Arg Val Ser Phe Trp Phe Val Val
             60                  65                  70
aca gac cct tca aaa aat cac acc ctt cct gct gtt gag gtg caa tca       340
Thr Asp Pro Ser Lys Asn His Thr Leu Pro Ala Val Glu Val Gln Ser
             75                  80                  85
gcc ata aga atg aac aag aac cgg atc aac aat gcc ttc ttt cta aat       388
Ala Ile Arg Met Asn Lys Asn Arg Ile Asn Asn Ala Phe Phe Leu Asn
90                  95                  100                 105
gac caa act ctg gaa ttt tta aaa atc cct tcc aca ctt gca cca ccc       436
Asp Gln Thr Leu Glu Phe Leu Lys Ile Pro Ser Thr Leu Ala Pro Pro
                 110                 115                 120
atg gac cca tct gtg ccc atc tgg att att ata ttt ggt gtg ata ttt       484
Met Asp Pro Ser Val Pro Ile Trp Ile Ile Ile Phe Gly Val Ile Phe
                 125                 130                 135
tgc atc atc ata gtt gca att gca cta ctg att tta tca ggg atc tgg       532
Cys Ile Ile Ile Val Ala Ile Ala Leu Leu Ile Leu Ser Gly Ile Trp
             140                 145                 150
caa cgt ada ara aag aac aaa gaa cca tct gaa gtg gat gac gct gaa       580
Gln Arg Xaa Xaa Lys Asn Lys Glu Pro Ser Glu Val Asp Asp Ala Glu
     155                 160                 165
rat aak tgt gaa aac atg atc aca att gaa aat ggc atc ccc tct gat       628
Xaa Xaa Cys Glu Asn Met Ile Thr Ile Glu Asn Gly Ile Pro Ser Asp
170                 175                 180                 185
ccc ctg gac atg aag gga ggg cat att aat gat gcc ttc atg aca gag       676
Pro Leu Asp Met Lys Gly Gly His Ile Asn Asp Ala Phe Met Thr Glu
                 190                 195                 200
gat gag agg ctc acc cct ctc tgaagggctg ttgttctgct tcctcaaraa         727
Asp Glu Arg Leu Thr Pro Leu
                 205
attaaacatt tgtttctgtg tgactgctga gcatcctgaa ataccaagag cagatcatat   787
wttttgtttc accattcttc ttttgtaata aatttgaat gtgcttgaaa aaaaaaaaaa     847
```

```
<210> 28
<211> 14
<212> PRT
<213> Homo Sapiens
<220>
<221> SIGNAL
<222> 1..14
<223> Von Heijne matrix
        score 10.7
        seq LWLLFFLVTAIHA/EL
<400> 28
Met Leu Trp Leu Leu Phe Phe Leu Val Thr Ala Ile His Ala
1               5                   10


<210> 29
<211> 25
<212> DNA
<213> Artificial Sequence
<220>
<223> oligonucleotide
<400> 29
gggaagatgg agatagtatt gcctg                                          25


<210> 30
<211> 26
<212> DNA
<213> Artificial Sequence
<220>
<223> oligonucleotide
<400> 30
ctgccatgta catgatagag agattc                                         26


<210> 31
<211> 546
<212> DNA
<213> Homo Sapiens
<220>
<221> promoter
<222> 1..517
<220>
<221> transcription start site
<222> 518
<220>
<221> protein_bind
<222> 17..25
<223> matinspector prediction
        name CMYB_01
        score 0.983
        sequence tgtcagttg
<220>
<221> protein_bind
<222> complement(18..27)
<223> matinspector prediction
        name MYOD_Q6
        score 0.961
        sequence cccaactgac
```
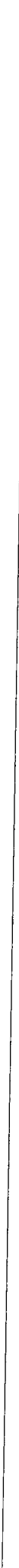
```
<220>
<221> protein_bind
<222> complement(75..85)
<223> matinspector prediction
       name S8_01
       score 0.960
       sequence aatagaattag
<220>
<221> protein_bind
<222> 94..104
<223> matinspector prediction
       name S8_01
       score 0.966
       sequence aactaaattag
<220>
<221> protein_bind
<222> complement(129..139)
<223> matinspector prediction
       name DELTAEF1_01
       score 0.960
       sequence gcacacctcag
<220>
<221> protein_bind
<222> complement(155..165)
<223> matinspector prediction
       name GATA_C
       score 0.964
       sequence agataaatcca
<220>
<221> protein_bind
<222> 170..178
<223> matinspector prediction
       name CMYB_01
       score 0.958
       sequence cttcagttg
<220>
<221> protein_bind
<222> 176..189
<223> matinspector prediction
       name GATA1_02
       score 0.959
       sequence ttgtagataggaca
<220>
<221> protein_bind
<222> 180..190
<223> matinspector prediction
       name GATA_C
       score 0.953
       sequence agataggacat
<220>
<221> protein_bind
<222> 284..299
<223> matinspector prediction
       name TAL1ALPHAE47_01
       score 0.973
       sequence cataacagatggtaag
<220>
<221> protein_bind
```

141
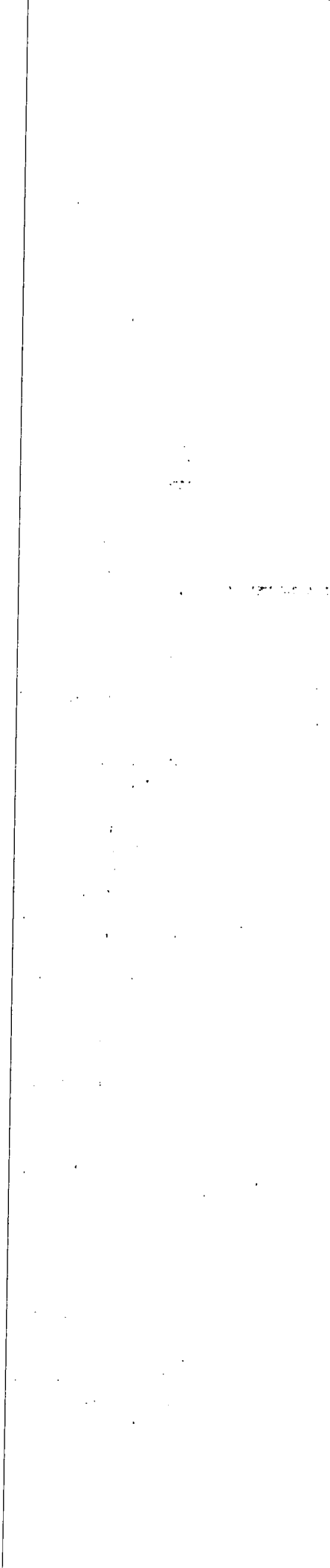
```
<222> 284..299
<223> matinspector prediction
      name TAL1BETAE47_01
      score 0.983
      sequence cataacagatggtaag
<220>
<221> protein_bind
<222> 284..299
<223> matinspector prediction
      name TAL1BETAITF2_01
      score 0.978
      sequence cataacagatggtaag
<220>
<221> protein_bind
<222> complement(287..296)
<223> matinspector prediction
      name MYOD_Q6
      score 0.954
      sequence accatctgtt
<220>
<221> protein_bind
<222> complement(302..314)
<223> matinspector prediction
      name GATA1_04
      score 0.953
      sequence tcaagataaagta
<220>
<221> protein_bind
<222> 393..405
<223> matinspector prediction
      name IK1_01
      score 0.963
      sequence agttgggaattcc
<220>
<221> protein_bind
<222> 393..404
<223> matinspector prediction
      name IK2_01
      score 0.985
      sequence agttgggaattc
<220>
<221> protein_bind
<222> 396..405
<223> matinspector prediction
      name CREL_01
      score 0.962
      sequence tgggaattcc
<220>
<221> protein_bind
<222> 423..436
<223> matinspector prediction
      name GATA1_02
      score 0.950
      sequence tcagtgatatggca
<220>
<221> protein_bind
<222> complement(478..489)
<223> matinspector prediction
```

13

```
        name SRY_02
        score 0.951
        sequence taaaacaaaaca
<220>
<221> protein_bind
<222> 486..493
<223> matinspector prediction
        name E2F_02
        score 0.957
        sequence tttagcgc
<220>
<221> protein_bind
<222> complement(514..521)
<223> matinspector prediction
        name MZF1_01
        score 0.975
        sequence tgagggga
<400> 31
tgagtgcagt gttacatgtc agttgggtta agtttgttaa tgtcattcaa atcttctatg      60
tcttgatttg cctgctaatt ctattatttc tggaactaaa ttagtttgat ggttctatta     120
gttattgact gaggtgtgct aatctcccat tatgtggatt tatctatttc ttcagttgta     180
gataggacat tgatagatac ataagtacca ggacaaaagc agggagatct tttttccaaa     240
atcaggagaa aaaaatgaca tctggaaaac ctatagggaa aggcataaca gatggtaagg     300
atactttatc ttgagtagga gagccttcct gtggcaacgt ggagaaggga agaggtcgta     360
gaattgagga gtcagctcag ttagaagcag ggagttggga attccgttca tgtgatttag     420
catcagtgat atggcaaatg tgggactaag ggtagtgatc agagggttaa aattgtgtgt     480
tttgtttttag cgctgctggg gcatcgcctt gggtcccctc aaacagattc ccatgaatct     540
cttcat                                                               546


<210> 32
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> oligonucleotide
<400> 32
gtaccaggga ctgtgaccat tgc                                              23


<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> oligonucleotide
<400> 33
ctgtgaccat tgctcccaag agag                                             24


<210> 34
<211> 861
<212> DNA
<213> Homo Sapiens
<220>
<221> promoter
<222> 1..806
<220>
<221> transcription start site
<222> 807
<220>
```

14

```
<221> protein_bind
<222> complement(60..70)
<223> matinspector prediction
       name NFY_Q6
       score 0.956
       sequence ggaccaatcat
<220>
<221> protein_bind
<222> 70..77
<223> matinspector prediction
       name MZF1_01
       score 0.962
       sequence cctgggga
<220>
<221> protein_bind
<222> 124..132
<223> matinspector prediction
       name CMYB_01
       score 0.994
       sequence tgaccgttg
<220>
<221> protein_bind
<222> complement(126..134)
<223> matinspector prediction
       name VMYB_02
       score 0.985
       sequence tccaacggt
<220>
<221> protein_bind
<222> 135..143
<223> matinspector prediction
       name STAT_01
       score 0.968
       sequence ttcctggaa
<220>
<221> protein_bind
<222> complement(135..143)
<223> matinspector prediction
       name STAT_01
       score 0.951
       sequence ttccaggaa
<220>
<221> protein_bind
<222> complement(252..259)
<223> matinspector prediction
       name MZF1_01
       score 0.956
       sequence ttggggga
<220>
<221> protein_bind
<222> 357..368
<223> matinspector prediction
       name IK2_01
       score 0.965
       sequence gaatgggatttc
<220>
<221> protein_bind
<222> 384..391
```

```
<223> matinspector prediction
      name MZF1_01
      score 0.986
      sequence agagggga
<220>
<221> protein_bind
<222> complement(410..421)
<223> matinspector prediction
      name SRY_02
      score 0.955
      sequence gaaaacaaaaca
<220>
<221> protein_bind
<222> 592..599
<223> matinspector prediction
      name MZF1_01
      score 0.960
      sequence gaagggga
<220>
<221> protein_bind
<222> 618..627
<223> matinspector prediction
      name MYOD_Q6
      score 0.981
      sequence agcatctgcc
<220>
<221> protein_bind
<222> 632..642
<223> matinspector prediction
      name DELTAEF1_01
      score 0.958
      sequence tcccaccttcc
<220>
<221> protein_bind
<222> complement(813..823)
<223> matinspector prediction
      name S8_01
      score 0.992
      sequence gaggcaattat
<220>
<221> protein_bind
<222> complement(824..831)
<223> matinspector prediction
      name MZF1_01
      score 0.986
      sequence agagggga
<220>
<221> misc_feature
<222> 335,376
<223> n=a, g, c or t
<400> 34
tactataggg cacgcgtggt cgacggccgg gctgttctgg agcagagggc atgtcagtaa      60
tgattggtcc ctggggaagg tctggctggc tccagcacag tgaggcattt aggtatctct     120
cggtgaccgt tggattcctg gaagcagtag ctgttctgtt tggatctggt agggacaggg     180
ctcagagggc taggcacgag ggaaggtcag aggagaaggs aggsarggcc cagtgagarg     240
ggagcatgcc ttcccccaac cctggcttsc ycttggymam agggcgktty tgggmacttr     300
aaytcagggc ccaascagaa scacaggccc aktcntggct smaagcacaa tagcctgaat     360
gggatttcag gttagncagg gtgagagggg aggctctctg gcttagtttt gttttgtttt     420
```

16

```
ccaaatcaag gtaacttgct cccttctgct acgggccttg gtcttggctt gtcctcaccc    480
agtcggaact ccctaccact ttcaggagag tggtttagg cccgtggggc tgttctgttc     540
caagcagtgt gagaacatgg ctggtagagg ctctagctgt gtgcggggcc tgaaggggag    600
tgggttctcg cccaaagagc atctgcccat ttcccacctt cccttctccc accagaagct    660
tgcctgagct gtttggacaa aaatccaaac cccacttggc tactctggcc tggcttcagc    720
ttggaaccca atacctaggc ttacaggcca tcctgagcca ggggcctctg gaaattctct    780
tcctgatggt cctttaggtt tgggcacaaa atataattgc ctctcccctc tcccattttc    840
tctcttggga gcaatggtca c                                             861


<210> 35
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> oligonucleotide
<400> 35
ctgggatgga aggcacggta                                                20


<210> 36
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> oligonucleotide
<400> 36
gagaccacac agctagacaa                                                20


<210> 37
<211> 555
<212> DNA
<213> Homo Sapiens
<220>
<221> promoter
<222> 1..500
<220>
<221> transcription start site
<222> 501
<220>
<221> protein_bind
<222> 191..206
<223> matinspector prediction
        name ARNT_01
        score 0.964
        sequence ggactcacgtgctgct
<220>
<221> protein_bind
<222> 193..204
<223> matinspector prediction
        name NMYC_01
        score 0.965
        sequence actcacgtgctg
<220>
<221> protein_bind
<222> 193..204
<223> matinspector prediction
        name USF_01
        score 0.985
        sequence actcacgtgctg
```

17 146

```
<220>
<221> protein_bind
<222> complement(193..204)
<223> matinspector prediction
        name USF_01
        score 0.985
        sequence cagcacgtgagt
<220>
<221> protein_bind
<222> complement(193..204)
<223> matinspector prediction
        name NMYC_01
        score 0.956
        sequence cagcacgtgagt
<220>
<221> protein_bind
<222> complement(193..204)
<223> matinspector prediction
        name MYCMAX_02
        score 0.972
        sequence cagcacgtgagt
<220>
<221> protein_bind
<222> 195..202
<223> matinspector prediction
        name USF_C
        score 0.997
        sequence tcacgtgc
<220>
<221> protein_bind
<222> complement(195..202)
<223> matinspector prediction
        name USF_C
        score 0.991
        sequence gcacgtga
<220>
<221> protein_bind
<222> complement(210..217)
<223> matinspector prediction
        name MZF1_01
        score 0.968
        sequence catgggga
<220>
<221> protein_bind
<222> 397..410
<223> matinspector prediction
        name ELK1_02
        score 0.963
        sequence ctctccggaagcct
<220>
<221> protein_bind
<222> 400..409
<223> matinspector prediction
        name CETS1P54_01
        score 0.974
        sequence tccggaagcc
<220>
<221> protein_bind
```

18
147

```
<222>  complement(460..470)
<223>  matinspector prediction
       name AP1_Q4
       score 0.963
       sequence agtgactgaac
<220>
<221>  protein_bind
<222>  complement(460..470)
<223>  matinspector prediction
       name AP1FJ_Q2
       score 0.961
       sequence agtgactgaac
<220>
<221>  protein_bind
<222>  547..555
<223>  matinspector prediction
       name PADS_C
       score 1.000
       sequence tgtggtctc
<400>  37
ctatagggca cgcktggtcg acggcccggg ctggtctggt ctgtkgtgga gtcgggttga     60
aggacagcat ttgtkacatc tggtctactg caccttccct ctgccgtgca cttggccttt    120
kawaagctca gcaccggtgc ccatcacagg gccggcagca cacacatccc attactcaga    180
aggaactgac ggactcacgt gctgctccgt ccccatgagc tcagtggacc tgtctatgta    240
gagcagtcag acagtgcctg ggatagagtg agagttcagc cagtaaatcc aagtgattgt    300
cattcctgtc tgcattagta actcccaacc tagatgtgaa aacttagttc tttctcatag    360
gttgctctgc ccatggtccc actgcagacc caggcactct ccggaagcct ggaaatcacc    420
cgtgtcttct gcctgctccc gctcacatcc cacacttgtg ttcagtcact gagttacaga    480
ttttgcctcc tcaatttctc ttgtcttagt cccatcctct gttcccctgg ccagtttgtc    540
tagctgtgtg gtctc                                                      555


<210>  38
<211>  19
<212>  DNA
<213>  Artificial Sequence
<220>
<223>  oligonucleotide
<400>  38
ggccatacac ttgagtgac                                                   19


<210>  39
<211>  19
<212>  DNA
<213>  Artificial Sequence
<220>
<223>  oligonucleotide
<400>  39
atatagacaa acgcacacc                                                   19


<210>  40
<211>  1098
<212>  DNA
<213>  Homo sapiens
<220>
<221>  sig_peptide
<222>  173..211
<223>  Von Heijne matrix
       score 4.19999980926514
```

19

```
        seq MLAVSLTVPLLGA/MM
<220>
<221> polyA_signal
<222> 1063..1068
<220>
<221> polyA_site
<222> 1087..1098
<220>
<221> misc_feature
<222> 144..467
<223> homology
      id :AA057573
      est
<220>
<221> misc_feature
<222> 510..640
<223> homology
      id :AA057573
      est
<220>
<221> misc_feature
<222> 436..523
<223> homology
      id :AA057573
      est
<220>
<221> misc_feature
<222> 708..786
<223> homology
      id :AA057573
      est
<220>
<221> misc_feature
<222> 635..682
<223> homology
      id :AA057573
      est
<220>
<221> misc_feature
<222> 625..1084
<223> homology
      id :N57409
      est
<220>
<221> misc_feature
<222> 779..1084
<223> homology
      id :R71351
      est
<220>
<221> misc_feature
<222> 144..506
<223> homology
      id :H12619
      est
<220>
<221> misc_feature
<222> 90..467
```

20

```
<223> homology
      id :T03538
      est
<220>
<221> misc_feature
<222> 314..523
<223> homology
      id :T34150
      est
<220>
<221> misc_feature
<222> 567..687
<223> homology
      id :T34150
      est
<220>
<221> misc_feature
<222> 686..730
<223> homology
      id :T34150
      est
<220>
<221> misc_feature
<222> 510..553
<223> homology
      id :T34150
      est
<220>
<221> misc_feature
<222> 550..579
<223> homology
      id :T34150
      est
<220>
<221> misc_feature
<222> 144..523
<223> homology
      id :N32314
      est
<220>
<221> misc_feature
<222> 510..553
<223> homology
      id :N32314
      est
<220>
<221> misc_feature
<222> 352..523
<223> homology
      id :T77966
      est
<220>
<221> misc_feature
<222> 218..351
<223> homology
      id :T77966
      est
<220>
```

21

150

```
agtgaggtgg tttctgcggg tgaggctggc gcccgtacca tgagcgaggc ggacgggctg      60
cgacagcgcc ggcccctgcg gcccgcaagt cgtcacagac gatgatggcc aggccccgga     120
ggctaaggac ggcagctcct ttagcggcag agttttccga gtgaccttct tg atg ctg     178
                                                           Met Leu

gct gtt tct ctc acc gtt ccc ctg ctt gga gcc atg atg ctg ctg gaa     226
Ala Val Ser Leu Thr Val Pro Leu Leu Gly Ala Met Met Leu Leu Glu
    -10                 -5                  1                   5
tct cct ata gat cca cag cct ctc agc ttc aaa gaa ccc ccg ctc ttg     274
Ser Pro Ile Asp Pro Gln Pro Leu Ser Phe Lys Glu Pro Pro Leu Leu
                10                  15                  20
ctt ggt gtt ctg cat cca aat acg aag ctg cga cag gca gaa agg ctg     322
Leu Gly Val Leu His Pro Asn Thr Lys Leu Arg Gln Ala Glu Arg Leu
            25                  30                  35
ttt gaa aat caa ctt gtt gga ccg gag tcc ata gca cat att ggg gat     370
Phe Glu Asn Gln Leu Val Gly Pro Glu Ser Ile Ala His Ile Gly Asp
        40                  45                  50
gtg atg ttt act ggg aca gca gat ggc cgg gtc gta aaa ctt gaa aat     418
Val Met Phe Thr Gly Thr Ala Asp Gly Arg Val Val Lys Leu Glu Asn
    55                  60                  65
ggt gaa ata gag acc att gcc cgg ttt ggt tcg ggc cct tgc aaa acc     466
Gly Glu Ile Glu Thr Ile Ala Arg Phe Gly Ser Gly Pro Cys Lys Thr
70                  75                  80                  85
cga ggt gat gag cct gtg tgt ggg aga ccc ctg ggt atc cgt ggc agg     514
Arg Gly Asp Glu Pro Val Cys Gly Arg Pro Leu Gly Ile Arg Gly Arg
                90                  95                  100
gcc caa tgg gac tct ctt tgt ggc cga tgc ata caa agg gac tat ttg     562
Ala Gln Trp Asp Ser Leu Cys Gly Arg Cys Ile Gln Arg Asp Tyr Leu
            105                 110                 115
aag taaatccctg gaaacgtgaa gtgaaactgc tgctgtcctc cgagacaccc     615
Lys
attgagggga agaacatgtc ctttgtgaat gatcttacag tcactcagga tgggaggaag     675
atttatttca ccgattctag cagcaaatgg caaagacgag actacctgct tctggtgatg     735
gagggcacag atgacgggcg cctgctggag tatgatactg tgaccaggga agtaaaagtt     795
ttattggacc agctgcggtt cccgaatgga gtccagctgt ctcctgcaga agactttgtc     855
ctggtggcag aaacaaccat ggccaggata cgaagagtct acgtttctgg cctgatgaag     915
ggcggggctg atctgtttgt ggagaacatg cctggatttc cagacaacat ccggcccagc     975
agctctgggg ggtactgggt gggcatgtcg accatccgcc ctaaccctgg tttttccatg    1035
ctggatttct tatctgagag accctggatt aaaaggatga tttttaangg taaaaaaaaa    1095
aaa                                                                  1098
```

<210> 41
<211> 855

<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 267..371
<223> Von Heijne matrix
        score 5.90000009536743
        seq LCGLLHLWLKVFS/LK
<220>
<221> polyA_signal
<222> 817..822
<220>
<221> polyA_site
<222> 842..855
<220>
<221> misc_feature
<222> 608..811
<223> homology
        id :M85769
        est
<400> 41
```
acaatcagtt tgccaatacc tcagaaacaa atacctcgga caaatctttc tctaaagacc      60
tcagtcagat actagtcaat atcaaatcat gtagatggcg gcattttagg cctcggacac     120
catccctaca tgacagtgac aatgatgaac tctcctgtag aaaattatat aggagtataa     180
accgaacagg aacagcacaa cctgggaccc agacatgcag tacctctacg caaagtaaaa     240
gtagcagtgg ttcagcacac tttggt atg ttg act gtt aat gat gta cgt ttc     293
                              Met Leu Thr Val Asn Asp Val Arg Phe
                              -35                 -30
tat aga aat gtc agg tcc aac cat ttc cca ttt gtt cga cta tgt ggt     341
Tyr Arg Asn Val Arg Ser Asn His Phe Pro Phe Val Arg Leu Cys Gly
    -25                 -20                 -15
ctg tta cat tta tgg ctt aaa gtc ttt tct ctt aaa cag tta aaa aaa     389
Leu Leu His Leu Trp Leu Lys Val Phe Ser Leu Lys Gln Leu Lys Lys
-10                 -5                  1                   5
aaa tct tgg tct aag tat tta ttt gaa tcc tgt tgc tat agg agt ttg     437
Lys Ser Trp Ser Lys Tyr Leu Phe Glu Ser Cys Cys Tyr Arg Ser Leu
        10                  15                  20
tat gtg tgt gtc ttc att taaacatacc tgcatacaaa gatggtttat           485
Tyr Val Cys Val Phe Ile
        25
ttctatttaa tatgtgacat ttgtttcctg gatatagtcc gtgaaccaca agatttatca     545
tattttttcaa taatatgaga agaaaatggg ccgtaaattg ttaaccattt tatgttcaga     605
tatttctcta gttttttacct agtttgcttt aacatagaga ccagcaagtg aatatatatg     665
cataacctta tatgttgaca caataattca gaataatttg ttaaagataa actaattttt     725
cagagaagaa catttaaagg gttaatattt ttgaaacgtt ttcagataat atctatttga     785
ttattgtggc ttctatttga aatgtgtcta aaataaaatg ctgtttattt aaaatgaaaa     845
aaaaaaaaaa                                                            855
```

<210> 42
<211> 1176
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 174..266
<223> Von Heijne matrix
        score 3.5
        seq WSPLSTRSGGTHA/CS

```
<220>
<221> polyA_signal
<222> 1144..1149
<220>
<221> polyA_site
<222> 1165..1176
<220>
<221> misc_feature
<222> 886..1134
<223> homology
      id :AA595193
      est
<220>
<221> misc_feature
<222> 756..894
<223> homology
      id :AA595193
      est
<220>
<221> misc_feature
<222> 655..755
<223> homology
      id :AA595193
      est
<220>
<221> misc_feature
<222> 167..367
<223> homology
      id :W81213
      est
<220>
<221> misc_feature
<222> 66..172
<223> homology
      id :W81213
      est
<220>
<221> misc_feature
<222> 429..508
<223> homology
      id :W81213
      est
<220>
<221> misc_feature
<222> 756..894
<223> homology
      id :AA150887
      est
<220>
<221> misc_feature
<222> 536..643
<223> homology
      id :AA150887
      est
<220>
<221> misc_feature
<222> 655..755
<223> homology
```

24

```
        id :AA150887
        est
<220>
<221> misc_feature
<222> 429..643
<223> homology
        id :AA493644
        est
<220>
<221> misc_feature
<222> 655..755
<223> homology
        id :AA493644
        est
<220>
<221> misc_feature
<222> 429..643
<223> homology
        id :AA493494
        est
<220>
<221> misc_feature
<222> 655..755
<223> homology
        id :AA493494
        est
<220>
<221> misc_feature
<222> 500..643
<223> homology
        id :AA179182
        est
<220>
<221> misc_feature
<222> 655..755
<223> homology
        id :AA179182
        est
<220>
<221> misc_feature
<222> 756..847
<223> homology
        id :AA179182
        est
<220>
<221> misc_feature
<222> 3..338
<223> homology
        id :HUM524F05B
        est
<220>
<221> misc_feature
<222> 334..374
<223> homology
        id :HUM524F05B
        est
<220>
<221> misc_feature
```

<222> 886..1134
<223> homology
     id :AA398156
     est
<220>
<221> misc_feature
<222> 756..894
<223> homology
     id :AA398156
     est
<220>
<221> misc_feature
<222> 590,601
<223> n=a, g, c or t
<400> 42

```
aaaaacaata ggacggaaac gccgaggaac ccggctgagg cggcagagca tcctggccag     60
aacaagccaa ggagccaaga cgagagggac acacggacaa acaacagaca gaagacgtac    120
tggccgctgg actccgctgc ctcccccatc tccccgccat ctgcgcccgg agg atg       176
                                                                 Met

agc cca gcc ttc agg gcc atg gat gtg gag ccc cgc gcc aaa ggc tcc     224
Ser Pro Ala Phe Arg Ala Met Asp Val Glu Pro Arg Ala Lys Gly Ser
-30                 -25                 -20                 -15

ttc tgg agc cct ttg tcc acc agg tcg ggg ggc act cat gcg tgc tcc     272
Phe Trp Ser Pro Leu Ser Thr Arg Ser Gly Gly Thr His Ala Cys Ser
                -10                 -5                  1

gct tca atg aga caa ccc tgg gca agc ccc tgg tcc caa ggg aac atc     320
Ala Ser Met Arg Gln Pro Trp Ala Ser Pro Trp Ser Gln Gly Asn Ile
         5                  10                  15

agt tct acg aga ccc tcc ctg ctg aga tgc gca aat tct ctc ccc agt     368
Ser Ser Thr Arg Pro Ser Leu Leu Arg Cys Ala Asn Ser Leu Pro Ser
    20                  25                  30

aca aag gac aaa gcc aaa ggc ccc ttg tta gct ggc cat ccc tgc ccc     416
Thr Lys Asp Lys Ala Lys Gly Pro Leu Leu Ala Gly His Pro Cys Pro
35                  40                  45                  50

att ttt tcc cct ggt cct ttc ccc tgt ggc cac agg gaa gtg tgg cct     464
Ile Phe Ser Pro Gly Pro Phe Pro Cys Gly His Arg Glu Val Trp Pro
                55                  60                  65

gaa tac ccc acc ccg gct cct ctg cac cca gag ctg ggg gcc acc tca     512
Glu Tyr Pro Thr Pro Ala Pro Leu His Pro Glu Leu Gly Ala Thr Ser
                70                  75                  80

gaa gtg tca tct ctc tct gag cac gsa ttc ccc tgc agc agt cga gga     560
Glu Val Ser Ser Leu Ser Glu His Xaa Phe Pro Cys Ser Ser Arg Gly
         85                  90                  95

ctg agc aga ttg agt gat gct ggg gca gan adg cct gag ang aaa ggt     608
Leu Ser Arg Leu Ser Asp Ala Gly Ala Xaa Xaa Pro Glu Xaa Lys Gly
    100                 105                 110

gtt cag cca gtc gtt tgt aag gcg ctc gkc ggm act gct gaa acg ccc     656
Val Gln Pro Val Val Cys Lys Ala Leu Xaa Gly Thr Ala Glu Thr Pro
115                 120                 125                 130

cca ccc tgacagcccc atcctcaaag actgtcttaa ttactcatgg caggttctag     712
Pro Pro

agacttaagg ggaaaagctg ctttcaaggc caccacatgt ctggtgctcc ccmaccagst    772
statctgcct wgtgttcatt ttgytatttt gtgasgtgag acagcaaaga ccaataaaaa    832
catattttat aagaacaaaa ggcytgggtg cctacccgkg tggggcacw gtgggaagcc    892
ttctgmtagg gtgtcttgtg ctgtrtggyt tgttttgttt gccccyttat tttgctttgc    952
ttacccagtc ttcccytamt yttggatgst tyttaaccct caggcaaacc tgtgttcccc   1012
ctgtattcag gstytgcttt aaagcaagcc atgaggctgt tggagtttct gtttagggca   1072
ttaaaaattc ccgcaaacta taaagagcaa tgttttcagt yttttaggat tagaagaatt   1132
```

26

acataaaaat taataaacat tttcaatgat ggaaaaaaaa aaaa   1176

<210> 43
<211> 648
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 460..555
<223> Von Heijne matrix
      score 4
      seq FSFMLLGMGGCLP/GF
<220>
<221> polyA_signal
<222> 614..619
<220>
<221> polyA_site
<222> 635..648
<400> 43
aattctggcc cagcttcttc cccagctcta tcctgcttcc ctccatctcc tataggattc   60
tccttagagt tctccctcca ttagtagttg tcttagggtc tgtttctggg gagccctgcc   120
taagactcat gctacaagaa gttaaataag tttcccgaag tcacacagct agcctctcat   180
ccctttttcta ctgagaggaa gtggaatgca ctccgacaag gataaggttt tattgtgagc   240
tggccttgga attaaaccac caccaacaca cttttggatt atcagaaggt ggaaggagtg   300
caaatgccag ttacggtgat gcgttcaaca tccttatttc cagtctttat gacgcctttc   360
ctgaatcaca ggtgcattgg ggtgcttcct cctccccagg actcccaccc aactttgtga   420
acacaaccca cttagaggag ttatctcagc acattatga atg ttg ggg acc acg   474
                                               Met Leu Gly Thr Thr
                                                                -30
ggc ctc ggg aca cag ggt cct tcc cag cag gct ctg ggc ttt ttc tcc   522
Gly Leu Gly Thr Gln Gly Pro Ser Gln Gln Ala Leu Gly Phe Phe Ser
        -25   -20   -15
ttt atg tta ctt gga atg ggc ggg tgc ctg cct gga ttc ctg cta cag   570
Phe Met Leu Leu Gly Met Gly Gly Cys Leu Pro Gly Phe Leu Leu Gln
    -10   -5   1   5
cct ccc aat cga tct cct act ttg cct gca tcc acc ttt gcc cat   615
Pro Pro Asn Arg Ser Pro Thr Leu Pro Ala Ser Thr Phe Ala His
                    10   15   20
taaagtcaat tctccaccca taaaaaaaaa aaa   648

<210> 44
<211> 1251
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 79..369
<223> Von Heijne matrix
      score 4
      seq RLPLVVSFIASSS/AN
<220>
<221> polyA_signal
<222> 1217..1222
<220>
<221> polyA_site
<222> 1240..1251
<220>
<221> misc_feature

27

```
<222> 2..423
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 463..520
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 418..467
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 159..636
<223> homology
      id :AA044187
      est
<220>
<221> misc_feature
<222> 629..684
<223> homology
      id :AA044187
      est
<220>
<221> misc_feature
<222> 5..453
<223> homology
      id :AA131958
      est
<220>
<221> misc_feature
<222> 446..494
<223> homology
      id :AA131958
      est
<220>
<221> misc_feature
<222> 14..343
<223> homology
      id :W95957
      est
<220>
<221> misc_feature
<222> 323..467
<223> homology
      id :W95957
      est
<220>
<221> misc_feature
<222> 463..494
<223> homology
      id :W95957
      est
```

```
<220>
<221> misc_feature
<222> 14..475
<223> homology
      id :W95790
      est
<220>
<221> misc_feature
<222> 410..876
<223> homology
      id :AA461134
      est
<220>
<221> misc_feature
<222> 974..1195
<223> homology
      id :AA595195
      est
<220>
<221> misc_feature
<222> 769..982
<223> homology
      id :AA595195
      est
<220>
<221> misc_feature
<222> 1208..1237
<223> homology
      id :AA595195
      est
<220>
<221> misc_feature
<222> 223..522
<223> homology
      id :AA041216
      est
<220>
<221> misc_feature
<222> 518..636
<223> homology
      id :AA041216
      est
<220>
<221> misc_feature
<222> 774..1127
<223> homology
      id :N94607
      est
<220>
<221> misc_feature
<222> 690..765
<223> homology
      id :N94607
      est
<220>
<221> misc_feature
<222> 833..1195
<223> homology
```

29

```
          id :AA076410
          est
<400> 44
aaagtgacag cggagagaac caggsagccc agaaacccca ggcgtggaga ttgatcctgc        60
gagagaaggg ggttcatc atg gcg gat gac cta aag cga ttc ttg tat aaa         111
                    Met Ala Asp Asp Leu Lys Arg Phe Leu Tyr Lys
                        -95                     -90
aag tta cca agt gtt gaa ggg ctc cat gcc att gtt gtg tca gat aga         159
Lys Leu Pro Ser Val Glu Gly Leu His Ala Ile Val Val Ser Asp Arg
    -85                 -80                     -75
gat gga gta cct gtt att aaa gtg gca aat gac aat gct cca gag cat         207
Asp Gly Val Pro Val Ile Lys Val Ala Asn Asp Asn Ala Pro Glu His
-70                 -65                     -60                     -55
gct ttg cga cct ggt ttc tta tcc act ttt gcc ctt gca aca gac caa         255
Ala Leu Arg Pro Gly Phe Leu Ser Thr Phe Ala Leu Ala Thr Asp Gln
                -50                 -45                     -40
gga agc aaa ctt gga ctt tcc aaa aat aaa agt atc atc tgt tac tat         303
Gly Ser Lys Leu Gly Leu Ser Lys Asn Lys Ser Ile Ile Cys Tyr Tyr
                -35                 -30                     -25
aac acc tac cag gtg gtt caa ttt aat cgt tta cct ttg gtg gtg agt         351
Asn Thr Tyr Gln Val Val Gln Phe Asn Arg Leu Pro Leu Val Val Ser
                -20                 -15                     -10
ttc ata gcc agc agc agt gcc aat aca gga cta att gtc agc cta gaa         399
Phe Ile Ala Ser Ser Ser Ala Asn Thr Gly Leu Ile Val Ser Leu Glu
    -5                  1                   5                       10
aag gag ctt gct cca ttg ttt gaa gaa ctg aga caa gtt gtg gaa att         447
Lys Glu Leu Ala Pro Leu Phe Glu Glu Leu Arg Gln Val Val Glu Ile
                15                  20                      25
tct taatctgaca gtggtttcag tgtgtacctt atcttcatta taacaacaca            500
Ser
atatcaatcc agcaatcttt agactacaat aatgctttta tccatgtgct caagaaaggg       560
cccccttttc caacttatac taaagaacta gcatatagat gtaatttata gatagatcag       620
ttgctatatt ttctggtgta aggtctttct tatttagtga gatctaggga taccacagaa       680
atggttcagt ctatcacagc tcccatggag ttagtctggt caccagatat ggatgagaga       740
ttctattcag tggattagaa tcaaactggt acattgatcc acttgagccg ttaagtgctg       800
ccaattgtac aatatgccca ggcttgcaga ataaagccaa ctttttattg tgaataataa       860
taaggacata tttttcttca gattatgttt tatttctttg cattgagtga ggtacataaa       920
atggcttggt aaaagtaata aaatcagtac aatcactaac tttcctttgt acatattatt       980
ttgcagtata gatgaatatt actaatcagt ttgattattc tcagagggtg ctgctcttta      1040
atgaaaatga aaattatagc taatgttttt tcctcaaact ctgctttctg taaccaatca      1100
gtgtttaat gtttgtgtgt tcttcataaa atttaaatac aattcgttat tctgtttcca       1160
atgttagtat gtatgtaaac atgatagtac agccattttt ttcatatgtg agtaaaaata      1220
aaatagtatt tttaaaagta aaaaaaaaaa a                                     1251

<210> 45
<211> 1524
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 160..231
<223> Von Heijne matrix
      score 5.69999980926514
      seq ILGLLGLLGTLVA/ML
<220>
<221> polyA_signal
<222> 1510..1515
<220>
```

```
<221> polyA_site
<222> 1506..1519
<220>
<221> misc_feature
<222> 1048..1504
<223> homology
      id :AA552647
      est
<220>
<221> misc_feature
<222> 597..846
<223> homology
      id :AA345449
      est
<220>
<221> misc_feature
<222> 39..93
<223> homology
      id :AA345449
      est
<220>
<221> misc_feature
<222> 113..149
<223> homology
      id :AA345449
      est
<220>
<221> misc_feature
<222> 98..400
<223> homology
      id :T86266
      est
<220>
<221> misc_feature
<222> 1210..1489
<223> homology
      id :T86158
      est
<220>
<221> misc_feature
<222> 954..983
<223> homology
      id :AA116709
      est
<400> 45
agctgcttgt ggccacccac agacacttgt aaggaggaga gaagtcagcc tggcagagag      60
actctgaaat gassgattag aggtgttcaa ggragcaaag agcttcagcc tgaagacaag     120
ggagcagtcc ctgaagacgc ttctactgag aggtctgcc atg gcc tct ctt ggc       174
                                           Met Ala Ser Leu Gly
                                                           -20
ctc caa ctt gtg ggc tac atc cta ggc ctt ctg ggg ctt ttg ggc aca     222
Leu Gln Leu Val Gly Tyr Ile Leu Gly Leu Leu Gly Leu Leu Gly Thr
            -15                 -10                 -5
ctg gtt gcc atg ctg ctc ccc agc tgg aaa aca agt tct tat gtc ggt     270
Leu Val Ala Met Leu Leu Pro Ser Trp Lys Thr Ser Ser Tyr Val Gly
                1                   5                  10
gcc agc att gtg aca gca gtt ggc ttc tcc aag ggc ctc tgg atg gaa     318
Ala Ser Ile Val Thr Ala Val Gly Phe Ser Lys Gly Leu Trp Met Glu
```

31

```
tgt gcc aca cac agc aca ggc atc acc cag tgt gac atc tat agc acc        366
Cys Ala Thr His Ser Thr Gly Ile Thr Gln Cys Asp Ile Tyr Ser Thr
30               35                   40                   45
ctt ctg ggc ctg ccc gct gac atc cak gct gcc cag gcc atg atg gtg        414
Leu Leu Gly Leu Pro Ala Asp Ile Xaa Ala Ala Gln Ala Met Met Val
              50                   55                   60
aca tcc agt gca atc tcc tcc ctg gcc tgc att atc tct gtg gtg ggc        462
Thr Ser Ser Ala Ile Ser Ser Leu Ala Cys Ile Ile Ser Val Val Gly
              65                   70                   75
atg ara tgc aca gtc ttc tgc cag gaa tcc cga gcc aaa gac aga gtg       ·510
Met Xaa Cys Thr Val Phe Cys Gln Glu Ser Arg Ala Lys Asp Arg Val
              80                   85                   90
gcg gta gca ggt gga gtc ttt ttc atc ctt gga ggc ctc ctg gga ttc        558
Ala Val Ala Gly Gly Val Phe Phe Ile Leu Gly Gly Leu Leu Gly Phe
    95                   100                  105
att cct gtt gcc tgg aat ctt cat ggg atc cta cgg gac ttc tac tca        606
Ile Pro Val Ala Trp Asn Leu His Gly Ile Leu Arg Asp Phe Tyr Ser
110                  115                  120                  125
cca ctg gtg cct gac agc atg aaa ttt gag att gga gag gct ctt tac        654
Pro Leu Val Pro Asp Ser Met Lys Phe Glu Ile Gly Glu Ala Leu Tyr
              130                  135                  140
ttg ggc att att tct tcc ctg ttc tcc ctg ata gct gga atc atc ctc        702
Leu Gly Ile Ile Ser Ser Leu Phe Ser Leu Ile Ala Gly Ile Ile Leu
              145                  150                  155
tgc ttt tcc tgc tca tcc cag aga aat cgc tcc aac tac tac gat gcc        750
Cys Phe Ser Cys Ser Ser Gln Arg Asn Arg Ser Asn Tyr Tyr Asp Ala
              160                  165                  170
tac caa gcc caa cct ctt gcc aca agg agc tct cca agg cct ggt caa        798
Tyr Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser Pro Arg Pro Gly Gln
    175                  180                  185
cct ccc aaa gtc aag agt gag ttc aat tcc tac agc ctg aca ggg tat        846
Pro Pro Lys Val Lys Ser Glu Phe Asn Ser Tyr Ser Leu Thr Gly Tyr
190                  195                  200                  205
gtg tgaagaacca ggggccagag ctggggggtg gctgggtctg tgaaaaacag             899
Val
tggacagcac cccgagggcc acaggtgagg gacactacca ctggatcgtg tcagaaggtg      959
ctgctgaggg tagactgact ttggccattg gattgagcaa aggcagaaat gggggctagt     1019
gtaacagcat gcaggttgaa ttgccaagga tgctcgccat gccagccttt ctgttttcct     1079
caccttgctg ctcccctgcc ctaagtcccc aaccctcaac ttgaaacccc attcccttaa     1139
gccaggamtc agaggatccc tytgccctck ggtttamctg ggactccatc cccaaaccca     1199
ctaatcacat cccactgact gaccctctgt gatcaaagac cctccctctg gctgaggttg     1259
gstyttagct cattgctggg gatgggaagg agaagcagtg gctttystgg gcattgctyt     1319
aacctamtty tcaagcttcc ctccaaagaa amtgattggc cctggaacct ccatcccact     1379
yttgttatga ctccacagtg tccagamtaa tttgtgcatg aactgaaata aaaccatcct     1439
acggtatyca gggaacagaa agcaggatgc aggatgggag gacaggaagg cagcctggga     1499
catttaaaaa aataaaaaaa aaaaa                                           1524
```

<210> 46
<211> 610
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 106..201
<223> Von Heijne matrix
    score 8.80000019073486
    seq VPMLLLIVGGSFG/LR

```
<220>
<221> polyA_signal
<222> 577..582
<220>
<221> polyA_site
<222> 598..610
<220>
<221> misc_feature
<222> 68..167
<223> homology
      id :AA531561
      est
<220>
<221> misc_feature
<222> 166..262
<223> homology
      id :AA531561
      est
<220>
<221> misc_feature
<222> 423..520
<223> homology
      id :AA531561
      est
<220>
<221> misc_feature
<222> 518..564
<223> homology
      id :AA531561
      est
<220>
<221> misc_feature
<222> 276..313
<223> homology
      id :AA531561
      est
<220>
<221> misc_feature
<222> 41..70
<223> homology
      id :AA531561
      est
<220>
<221> misc_feature
<222> 41..262
<223> homology
      id :AA535454
      est
<220>
<221> misc_feature
<222> 423..520
<223> homology
      id :AA535454
      est
<220>
<221> misc_feature
<222> 518..564
<223> homology
```

33

```
            id :AA535454
      est
<220>
<221> misc_feature
<222> 276..313
<223> homology
            id :AA535454
      est
<220>
<221> misc_feature
<222> 46..262
<223> homology
            id :H81225
      est
<220>
<221> misc_feature
<222> 2..39
<223> homology
            id :H81225
      est
<220>
<221> misc_feature
<222> 455..493
<223> homology
            id :H81225
      est
<220>
<221> misc_feature
<222> 276..313
<223> homology
            'id :H81225
      est
<220>
<221> misc_feature
<222> 423..458
<223> homology
            id :H81225
      est
<220>
<221> misc_feature
<222> 53..262
<223> homology
            id :AA044291
      est
<220>
<221> misc_feature
<222> 423..520
<223> homology
            id :AA044291
      est
<220>
<221> misc_feature
<222> 518..564
<223> homology
            id :AA044291
      est
<220>
<221> misc_feature
```

34   163

```
<222> 276..313
<223> homology
      id :AA044291
      est
<220>
<221> misc_feature
<222> 125..262
<223> homology
      id :W47031
      est
<400> 46
aaagtgagtt aaggacgtac tcgtcttggt gagagcgtga stgctgagat ttgggagtct      60
gcgctaggcc cgcttggagt tctgagccga tggaagagtt cactc atg ttt gca ccc     117
                                                   Met Phe Ala Pro
                                                             -30
gcg gtg atg cgt gct ttt cgc aag aac aag act ctc ggc tat gga gtc      165
Ala Val Met Arg Ala Phe Arg Lys Asn Lys Thr Leu Gly Tyr Gly Val
            -25                     -20                     -15
ccc atg ttg ttg ctg att gtt gga ggt tct ttt ggt ctt cgt gag ttt      213
Pro Met Leu Leu Leu Ile Val Gly Gly Ser Phe Gly Leu Arg Glu Phe
            -10                     -5                      1
tct caa atc cga tat gat gct gtg aag agt aaa atg gat cct gag ctt      261
Ser Gln Ile Arg Tyr Asp Ala Val Lys Ser Lys Met Asp Pro Glu Leu
5                     10                      15                      20
gaa aaa aaa ccg aaa gag aat aaa ata tct tta gag tcg gaa tat gag      309
Glu Lys Lys Pro Lys Glu Asn Lys Ile Ser Leu Glu Ser Glu Tyr Glu
                25                      30                      35
gga agt atc tgt tgaagggcta ctatctttcc ttggcccttc tcccttgttg          361
Gly Ser Ile Cys
                40
ggactcaatc tccagactat ctccccagag aatcttgtca aggcttggct ttaagctttg     421
ttgggaaaat caaagactcc aagtttgatg actggaagaa tattcgagga cccaggcctt     481
gggaagatcc tgacctcctc caaggaagaa atccaggaaa gccttaagac taagacaact     541
tgactctgct gattcttttt tccttttttt ttttaaataa aaatactatt aactggaaaa     601
aaaaaaaaa                                                             610


<210> 47
<211> 1370
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 359..466
<223> Von Heijne matrix
      score 7.80000019073486
      seq LTFLFLHLPPSTS/LF
<220>
<221> polyA_signal
<222> 1334..1339
<220>
<221> polyA_site
<222> 1357..1370
<220>
<221> misc_feature
<222> 113..420
<223> homology
      id :R79290
      est
```

35 164

```
<220>
<221> misc_feature
<222> 406..482
<223> homology
        id :R79290
        est
<220>
<221> misc_feature
<222> 199..420
<223> homology
        id :R81173
        est
<220>
<221> misc_feature
<222> 406..514
<223> homology
        id :R81173
        est
<220>
<221> misc_feature
<222> 2..269
<223> homology
        id :R81277
        est
<220>
<221> misc_feature
<222> 406..646
<223> homology
        id :R74123
        est
<220>
<221> misc_feature
<222> 647..682
<223> homology
        id :R74123
        est
<220>
<221> misc_feature
<222> 439..646
<223> homology
        id :AA450228
        est
<220>
<221> misc_feature
<222> 647..739
<223> homology
        id :AA450228
        est
<220>
<221> misc_feature
<222> 406..646
<223> homology
        id :R02473
        est
<220>
<221> misc_feature
<222> 406..604
<223> homology
```

36

```
        id :T71107
      est
<220>
<221> misc_feature
<222> 71..282
<223> homology
        id :C06030
      est
<220>
<221> misc_feature
<222> 319..365
<223> homology
   .    id :C06030
      est
<220>
<221> misc_feature
<222> 2..57
<223> homology
        id :C06030
      est
<220>
<221> misc_feature
<222> 1173..1277
<223> homology
        id :N54909
      est
<220>
<221> misc_feature
<222> 1080..1177
<223> homology
        id :N54909
      est
<220>
<221> misc_feature
<222> 1273..1356
<223> homology
        id :N54909
      est
<220>
<221> misc_feature
<222> 1173..1277
<223> homology
        id :AA196824
      est
<220>
<221> misc_feature
<222> 1080..1177
<223> homology
        id :AA196824
      est
<220>
<221> misc_feature
<222> 1273..1356
<223> homology
        id :AA196824
      est
<400> 47
acaaggcaga gcttctgaat ttcaggcctt cattccagag ccctcttgtg gccaggcctt      60
```

```
cctttgctgg aggaaggtac acagggtgaa gctgawgstg tacttggggg atctccttgg    120
cctgttccac caagtgagag aaggtactta ctcttgtacc tcctgttcag ccaggtgcat    180
taacagacct ccctacagct gtaggaacta ctgtcccaga gctgaggcaa ggggatttct    240
caggtcattt ggagaacaag tgctttagta gtagtttaaa gtagtaactg ctactgtatt    300
tagtggggtg gaattcagaa gaaatttgaa gaccagatca tgggtggtct gcatgtga     358
atg aac ach ttt gag cca gac agc ctg gct gtc att gct ttc ttc ctc    406
Met Asn Thr Phe Glu Pro Asp Ser Leu Ala Val Ile Ala Phe Phe Leu
    -35                      -30                     -25
ccc att tgg acc ttc tct gcc ctt aca ttt ttg ttt ctc cat cta cca    454
Pro Ile Trp Thr Phe Ser Ala Leu Thr Phe Leu Phe Leu His Leu Pro
    -20                -15                  -10                   -5
cca tcc acc agt cta ttt att aac tta gca aga gga caa ata aag ggc    502
Pro Ser Thr Ser Leu Phe Ile Asn Leu Ala Arg Gly Gln Ile Lys Gly
                1                  5                  10
cct ctt ggc ttg att ttg ctt ctt tct ttc tgt gga gga tat act aag    550
Pro Leu Gly Leu Ile Leu Leu Leu Ser Phe Cys Gly Gly Tyr Thr Lys
        15                      20                  25
tgc gac ttt gcc cta tcc tat ttg gaa atc cct aac aga att gag ttt    598
Cys Asp Phe Ala Leu Ser Tyr Leu Glu Ile Pro Asn Arg Ile Glu Phe
    30                      35                  40
tct att atg gat cca aaa aga aaa aca aaa tgc taatgaagcc atcasgtcaa    651
Ser Ile Met Asp Pro Lys Arg Lys Thr Lys Cys
45                      50                  55
gggtcacatg ccaataaaca ataaattttc cagaagaaat gaaatccaac tagacaaata    711
aagtagagct tatgaaatgg ttcagtaagg atgagcttgt tgttttttgt tttgtttttgt    771
tttgtttttt taaagacgga gtctcgctct gtcactcagg ctggagtgca gtggtatgat    831
cttggctcac tgtaacctcc gcctcccggg ttcaagccat tctcctgcct cagtctcctg    891
agtagctggg attgcaggtg cgtgccacca tgcctggcta attttgtgt ttttggtaga     951
gacagggttt caccacgttg gtcgggctgg tctcgggctc ctgacctctt gatccgcctg    1011
ccttggcctc ccaaagtgat gggattacag atgtgagcca ccgtgcctag ccaaggatga    1071
gatttttaaa gtatgttcca gttctgtgtc atggttggaa gacagagtag gaaggatatg    1131
gaaaaggtca tggggaagca gaggtgattc atggctctgt ggaatttgag gtgaatggtt    1191
ccttattgtc taggccactt gtgaagaata tgagtcagtt attgccagcc ttggaattta    1251
cttctctagc ttacaatgga ccttttgaa ctgggaaaca ccttgtctgc attcacttta     1311
aaatgtcaaa actaattttt ataataaatg tttattttca catygaaaaa aaaaaaaaa     1370
```

<210> 48
<211> 791
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 191..286
<223> Von Heijne matrix
      score 8.80000019073486
      seq VPMLLLIVGGSFG/LR
<220>
<221> polyA_signal
<222> 755..760
<220>
<221> polyA_site
<222> 780..791
<220>
<221> misc_feature
<222> 361..531
<223> homology
      id :W73841
      est

38

```
<220>
<221> misc_feature
<222> 210..347
<223> homology
        id :W73841
        est
<220>
<221> misc_feature
<222> 548..637
<223> homology
        id :W73841
        est
<220>
<221> misc_feature
<222> 181..210
<223> homology
        id :W73841
        est
<220>
<221> misc_feature
<222> 361..530
<223> homology
        id :HSU74317
        est
<220>
<221> misc_feature
<222> 238..347
<223> homology
        id :HSU74317
        est
<220>
<221> misc_feature
<222> 568..637
<223> homology
        id :HSU74317
        est
<220>
<221> misc_feature
<222> 698..733
<223> homology
        id :HSU74317
        est
<220>
<221> misc_feature
<222> 361..531
<223> homology
        id :W47031
        est
<220>
<221> misc_feature
<222> 210..347
<223> homology
        id :W47031
        est
<220>
<221> misc_feature
<222> 148..210
<223> homology
```

```
        id :W47031
        est
<220>
<221> misc_feature
<222> 548..600
<223> homology
        id :W47031
        est
<220>
<221> misc_feature
<222> 129..347
<223> homology
        id :AA044118
        est
<220>
<221> misc_feature
<222> 437..531
<223> homology
        id :AA044118
        est
<220>
<221> misc_feature
<222> 361..454
<223> homology
        id :AA044118
        est
<220>
<221> misc_feature.
<222> 176..347
<223> homology
        id :AA293342
        est
<220>
<221> misc_feature
<222> 361..531
<223> homology
        id :AA293342
        est
<220>
<221> misc_feature
<222> 548..605
<223> homology
        id :AA293342
        est
<220>
<221> misc_feature
<222> 361..531
<223> homology
        id :AA531561
        est
<220>
<221> misc_feature
<222> 153..252
<223> homology
        id :AA531561
        est
<220>
<221> misc_feature
```

169

<222> 750
<223> n=a, g, c or t
<400> 48
aacaagtatg ttacgatggc tcgattgctt ttgcctagcg gaaaccattc actaaggacc   60
gagcaccaaa taaccaagga aaaggaagtg agttaaggac gtactcgtct tggtgagagc   120
gtgagctgct gagatttggg agtctgcgct aggcccgctt ggagttctga gccgatggaa   180
gagttcactc atg ttt gca ccc gcg gtg atg cgt gct ttt cgc aag aac   229
          Met Phe Ala Pro Ala Val Met Arg Ala Phe Arg Lys Asn
                   -30   -25   -20
aag act ctc ggc tat gga gtc ccc atg ttg ttg ctg att gtt gga ggt   277
Lys Thr Leu Gly Tyr Gly Val Pro Met Leu Leu Leu Ile Val Gly Gly
             -15   -10   -5
tct ttt ggt ctt cgt gag ttt tct caa atc cga tat gat gct gtg aag   325
Ser Phe Gly Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp Ala Val Lys
         1   5   10
ggt aaa atg gat cct gag ctt gaa aaa aaa ctg aaa gag aat aaa ata   373
Gly Lys Met Asp Pro Glu Leu Glu Lys Lys Leu Lys Glu Asn Lys Ile
     15   20   25
tct tta gag tcg gaa tat gag aaa atc aaa gac tcc aag ttt gat gac   421
Ser Leu Glu Ser Glu Tyr Glu Lys Ile Lys Asp Ser Lys Phe Asp Asp
30   35   40   45
tgg aag aat att cga gga ccc agg cct tgg gaa gat cct gac ctc ctc   469
Trp Lys Asn Ile Arg Gly Pro Arg Pro Trp Glu Asp Pro Asp Leu Leu
             50   55   60
caa gga aga aat cca gaa agc ctt aag act aag aca act tgactctgct   518
Gln Gly Arg Asn Pro Glu Ser Leu Lys Thr Lys Thr Thr
             65   70
gattctcttt tccttttttt ttttaaataa aaatactatt aactggactt cctaatatat   578
acttctatca agtggaaagg aaattccagg cccatggaaa cttggatatg ggtaatttgg   638
atggacaaaa ktaatctktc actaaaggtc atgtaccagg tttttatact tcccagctaa   698
ttccatctgt ggatgaaagt tgcaatgttg gcccccgtat kattttacac cntcgaaata   758
aaaaatgtga ataactgctc caaaaaaaaa aaa   791

<210> 49
<211> 1433
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 346..408
<223> Von Heijne matrix
      score 5.5
      seq SFLPSALVIWTSA/AF
<220>
<221> polyA_signal
<222> 1400..1405
<220>
<221> polyA_site
<222> 1420..1433
<220>
<221> misc_feature
<222> 268..634
<223> homology
      id :W02860
      est
<220>
<221> misc_feature
<222> 118..564

41

```
<223> homology
      id :N27248
      est
<220>
<221> misc_feature
<222> 268..697
<223> homology
      id :N44490
     .est
<220>
<221> misc_feature
<222> 582..687
<223> homology
      id :AA274731
      est
<220>
<221> misc_feature
<222> 65..369
<223> homology
      id :H94779
      est
<220>
<221> misc_feature
<222> 471..519
<223> homology
      id :H94779
      est
<220>
<221> misc_feature
<222> 61..399
<223> homology
      id :H09880
      est
<220>
<221> misc_feature
<222> 408..452
<223> homology
      id :H09880
      est
<220>
<221> misc_feature
<222> 484..699
<223> homology
      id :H04537
      est
<220>
<221> misc_feature
<222> 685..772
<223> homology
      id :H04537
      est
<220>
<221> misc_feature
<222> 454..486
<223> homology
      id :H04537
      est
<220>
```

```
<221> misc_feature
<222> 410..439
<223> homology
      id :H04537
      est
<220>
<221> misc_feature
<222> 572..687
<223> homology
      id :AA466632
      est
<220>
<221> misc_feature
<222> 260..444
<223> homology
      id :AA459511
      est
<220>
<221> misc_feature
<222> 449..567
<223> homology
      id :AA459511
      est
<220>
<221> misc_feature
<222> 117..184
<223> homology
      id :AA459511
      est
<220>
<221> misc_feature
<222> 260..464
<223> homology
      id :H57434
      est
<220>
<221> misc_feature
<222> 118..184
<223> homology
      id :H57434
      est
<220>
<221> misc_feature
<222> 56..113
<223> homology
      id :H57434
      est
<220>
<221> misc_feature
<222> 454..485
<223> homology
      id :H57434
      est
<220>
<221> misc_feature
<222> 115,1177,1181,1200,1210,1245,1254,1263
<223> n=a, g, c or t
<400> 49
```

43

```
actcctttta gcatagggge ttcggcgcca gcggccagcg ctagtcggtc tggtaagtgc          60
ctgatgccga gttccgtctc tcgcgtcttt tcctggtccc aggcaaagcg gasgnagatc         120
ctcaaacggc ctagtgcttc gcgcttccgg agaaaatcag cggtctaatt aattcctctg         180
gtttgttgaa gcagttacca agaatcttca acccttttccc acaaaagcta attgagtaca         240
cgttcctgtt gagtacacgt tcctgttgat ttacaaaagg tgcaggtatg agcaggtctg         300
aagactaaca ttttgtgaag ttgtaaaaca gaaaacctgt tagaa atg tgg tgg ttt        357
                                                    Met Trp Trp Phe
                                                        -20

cag caa ggc ctc agt ttc ctt cct tca gcc ctt gta att tgg aca tct        405
Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr Ser
        -15             -10                 -5

gct gct ttc ata ttt tca tac att act gca gta aca ctc cac cat ata        453
Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His Ile
    1               5               10                  .15

gac ccg gct tta cct tat atc agt gac act ggt aca gta gct cca gaa        501
Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro Glu
                20                  25                  30

aaa tgc tta ttt ggg gca atg cta aat att gcg gca gtt tta tgc att        549
Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys Ile
            35                  40                  45

gct acc att tat gtt cgt tat aag caa gtt cat gct ctg agt cct gaa        597
Ala Thr Ile Tyr Val Arg Tyr Lys Gln Val His Ala Leu Ser Pro Glu
            50                  55                  60

gag aac gtt atc atc aaa tta aac aag gct ggc ctt gta ctt gga ata        645
Glu Asn Val Ile Ile Lys Leu Asn Lys Ala Gly Leu Val Leu Gly Ile
    65                  70                  75

ctg agt tgt tta gga ctt tct att gtg gca aac ttc cag gaa aac aac        693
Leu Ser Cys Leu Gly Leu Ser Ile Val Ala Asn Phe Gln Glu Asn Asn
80                  85                  90                  95

cct ttt tgc tgc aca tgt aag tgg agc tgt gct tac ctt tgg tat ggg        741
Pro Phe Cys Cys Thr Cys Lys Trp Ser Cys Ala Tyr Leu Trp Tyr Gly
                100                 105                 110

ctc att ata tat gtt tgt tca gac cat cct ttc cta cca aaa tgc agc        789
Leu Ile Ile Tyr Val Cys Ser Asp His Pro Phe Leu Pro Lys Cys Ser
            115                 120                 125

cca aaa tcc aat ggc aaa aca agt ctt ctg gat cag act gtt gtt ggt        837
Pro Lys Ser Asn Gly Lys Thr Ser Leu Leu Asp Gln Thr Val Val Gly
            130                 135                 140

tat ctg gtg tgg agt aag tgc act tagcatgctg acttgctcat cagttttgca        891
Tyr Leu Val Trp Ser Lys Cys Thr
            145                 150

cagtggcaat tttgggactg atttagaaca gaaactccat tggaaccccg aggacaaagg        951
ttatgcgctt cacatgatca ctactgcagc agaatggtct atgtcatttt ccttctttgg       1011
ttttttcctg acttacattc gtgattttca gaaaatttcc ttacgggtgg aagccaactt       1071
acatggatta accctctatg acactgcacc ttgccctatt aacaatgaac gaacacggct       1131
actttccags aagatattag atgaaaggat aaaatatttc tgtaantgan ttastgastt       1191
ctcagggant tggggaaang gttcacagaa gttgcttavt tcttcatcrt gaanattttc       1251
aanccactta antcaaggct gacagstaac acgtgatgaa tgctgataat caggaaacat       1311
gaaagaagcc atttgcatag.attattytaa aggatatcat caagaagamt attaaaaaca       1371
cctatgccta tactttttta tytcagaaaa taaagtcaaa agactatgaa aaaaaaaaaa       1431
aa                                                                    1433

<210> 50
<211> 1158
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
```

<222> 214..339
<223> Von Heijne matrix
score 6.09999990463257
seq AILLLQSQCAYWA/LP
<220>
<221> polyA_signal
<222> 1133..1138
<220>
<221> polyA_site
<222> 1146..1158
<220>
<221> misc_feature
<222> 840..968
<223> homology
id :H64717
est
<220>
<221> misc_feature
<222> 858..968
<223> homology
id :H65208
est
<220>
<221> misc_feature
<222> 652
<223> n=a, g, c or t
<400> 50

```
aarttgagct tggggactgc agctgtgggg agatttcagt gcattgcctc ccctgggtgc      60
tcttcatctt ggatttgaaa gttgagagca gcatgttttg cccactgaaa ctcatcctgs     120
tgrsagtgta mtggattatt ccttgggcct gaatgacttg aatgtttccc cgcctgagct     180
aacagtccat gtgggtgatt cagctctgat ggg atg tgt ttt cca gag cac aga      234
                                    Met Cys Phe Pro Glu His Arg
                                                            -40
aga caa atg tat att caa gat aga ctg gac tct gtc acc agg aga gca       282
Arg Gln Met Tyr Ile Gln Asp Arg Leu Asp Ser Val Thr Arg Arg Ala
-35                     -30                 -25                 -20
cgc caa gga cga ata tgt gct ata cta tta ctc caa tct cag tgt gcc       330
Arg Gln Gly Arg Ile Cys Ala Ile Leu Leu Leu Gln Ser Gln Cys Ala
                -15                 -10                 -5
tat tgg gcg ctt cca gaa ccg cgt aca ctt gat ggg gga cat ctt atg       378
Tyr Trp Ala Leu Pro Glu Pro Arg Thr Leu Asp Gly Gly His Leu Met
            1               5                   10
caa tgatggctct ctcctgctcc aagatgtgca agaggctgac cagggaacct            431
Gln
atatctgtga atccgcctc aaaggggaga gccaggtgtt caagaaggcg gtggtactgc      491
atgtgcttcc agaggagccc aaaggtacgc aaatgcttac ttaaagaggg gccaaggggc     551
aagagctttc atgtgcaaga ggcaaggaaa ctgattatct tgagtaaatg ccagcctttg     611
ggctaagtac ttaccacaga gtgaatcttc aaagaaatga ntcattaaat tatttcagrt     671
cagaataaaa atakgagtta ttttagttaa kaataaaata ttgataatta ttgtattatt     731
actttaaaca cacttccccc tcacaaaagc cctgtgaagg atgttttgtt cacatataat     791
gtccaaatat gttttggaca catatttatt aaatggaata aatagtamtt gaaccctggc     851
accthtgaca acaaagtcya tgttyttttt actatgccct aataccttts atcagttatc     911
cacattgatg ctacatytgt attttatagg tacccatgt taggtgtttt gggggataga      971
aaagaaataa gcagkycagg ctcagtggct catgcctgta atcctagcat tttgggaggc    1031
tgaggcagca gaamtgcctg agccccaggg ttcaagactg cagtgagcta tgawggcacc    1091
actgcattyt agcctgggwg acagagcaag actytgttta aaataaaaaa agagaaaaaa    1151
aaaaaaa                                                              1158
```

45 174

```
<210> 51
<211> 850
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 372..437
<223> Von Heijne matrix
      score 6.09999990463257
      seq LFLTCLFWPLAAL/NV
<220>
<221> polyA_signal
<222> 812..817
<220>
<221> polyA_site
<222> 838..850
<220>
<221> misc_feature
<222> 128..424
<223> homology
      id :N78012
      est
<220>
<221> misc_feature
<222> 61..128
<223> homology
      id :N78012
      est
<220>
<221> misc_feature
<222> 483..554
<223> homology
      id :N78012
      est
<220>
<221> misc_feature
<222> 417..464
<223> homology
      id :N78012
      est
<220>
<221> misc_feature
<222> 460..500
<223> homology
      id :N78012
      est
<220>
<221> misc_feature
<222> 577..612
<223> homology
      id :N78012
      est
<220>
<221> misc_feature
<222> 612..649
<223> homology
      id :N78012
      est
```

175

```
<220>
<221> misc_feature
<222> 546..577
<223> homology
       id :N78012
       est
<220>
<221> misc_feature
<222> 29..63
<223> homology
       id :N78012
       est
<220>
<221> misc_feature
<222> 128..294
<223> homology
       id :W37233
       est
<220>
<221> misc_feature
<222> 370..509
<223> homology
       id :W37233
       est
<220>
<221> misc_feature
<222> 505..591
<223> homology
       id :W37233
       est
<220>
<221> misc_feature
<222> 293..330
<223> homology
       id :W37233
       est
<220>
<221> misc_feature
<222> 22..57
<223> homology
       id :W37233
       est
<220>
<221> misc_feature
<222> 95..128
<223> homology
       id :W37233
       est
<220>
<221> misc_feature
<222> 128..326
<223> homology
       id :AA186399
       est
<220>
<221> misc_feature
<222> 418..605
<223> homology
```

176

```
          id :AA186399
       est
<220>
<221> misc_feature
<222> 326..423
<223> homology
          id :AA186399
       est
<220>
<221> misc_feature
<222> 39..128
<223> homology
          id :AA186399
       est
<220>
<221> misc_feature
<222> 206..640
<223> homology
    .     id :W52489
       est
<400> 51
agacacttcc tggtgggatc cgagtgaggc gacggggtag gggttggcgc tcaggcggcg      60
accatggcgt atcacggcct cactgtgcct ctcattgtga tgagcgtgtt ctggggcttc     120
gtcggctttc ttggtgcctt ggttcatccc taagggtcct aaccggggag ttatcattac     180
catgttggtg acctgttcag tttgctgcta tctcttttgg ctgattgcaa ttctggccca     240
actcaaccct ctctttggac cgcaattgaa aaatgaaacc atctggtatc tgaagtatca     300
ttggccttga ggaagaagac atgctctaca gtgctcagtc tttgaggtca cgagaagaga     360
atgccttcta g atg caa aat cac ctc caa acc aga cca ctt ttc ttg act     410
            Met Gln Asn His Leu Gln Thr Arg Pro Leu Phe Leu Thr
                 -20               -15                  -10
tgc ctg ttt tgg cca tta gct gcc tta aac gtt aac agc aca ttt gaa     458
Cys Leu Phe Trp Pro Leu Ala Ala Leu Asn Val Asn Ser Thr Phe Glu
          -5                    1               5
tgc ctt att cta caa tgc agc gtg ttt tcc ttt gcc ttt ttt gca ctt     506
Cys Leu Ile Leu Gln Cys Ser Val Phe Ser Phe Ala Phe Phe Ala Leu
       10                 15                  20
tgg tgaattacgt gcctccataa cctgaactgt gccgactcca caaaacgatt          559
Trp
atgtactctt ctgagataga agatgctgtt cttctgagag atacgttact ctctccttgg     619
aatctgtgga tttgaaaatg gctcctgcct tctcacgtgg gaatcagtga agtgtttaga     679
aactgctgca agacaaacaa gactccagtg gggtggtcag taggaaaaca cgttcagagg     739
gaagaaccat ctcaacagaa tcgcaccaaa ctatactttc aggatgaatt cttctttct     799
gccatctttt ggaataaata ttttcctcct ttytatgtaa aaaaaaaaaa a           850


<210> 52
<211> 1107
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 132..215
<223> Von Heijne matrix
       score 3.59999990463257
    .. seq PLSDSWALLPASA/GV
<220>
<221> polyA_signal
<222> 1069..1074
<220>
```

```
<221> polyA_site
<222> 1094..1107
<220>
<221> misc_feature
<222> 177..392
<223> homology
      id :W80978
      est
<220>
<221> misc_feature
<222> 425..542
<223> homology
      id :W80978
      est
<220>
<221> misc_feature
<222> 43..114
<223> homology
      id :W80978
      est
<220>
<221> misc_feature
<222> 387..441
<223> homology
      id :W80978
      est
<220>
<221> misc_feature
<222> 113..165
<223> homology
      id :W80978
      est
<220>
<221> misc_feature
<222> 551..590
<223> homology
      id :W80978
      est
<220>
<221> misc_feature
<222> 166..314
<223> homology
      id :AA043154
      est
<220>
<221> misc_feature
<222> 27..181
<223> homology
      id :AA043154
      est
<220>
<221> misc_feature
<222> 425..564
<223> homology
      id :AA043154
      est
<220>
<221> misc_feature
```

49

```
<222> 387..441
<223> homology
      id :AA043154
      est
<220>
<221> misc_feature
<222> 309..352
<223> homology
      id :AA043154
      est
<220>
<221> misc_feature
<222> 549..580
<223> homology
      id :AA043154
      est
<220>
<221> misc_feature
<222> 601..1071
<223> homology
      id :AA126732
      est
<220>
<221> misc_feature
<222> 576..605
<223> homology
      id :AA126732
      est
<220>
<221> misc_feature
<222> 387..477
<223> homology
      id :AA161280
      est
<220>
<221> misc_feature
<222> 292..362
<223> homology
      id :AA161280
      est
<220>
<221> misc_feature
<222> 46..113
<223> homology
      id :AA161280
      est
<220>
<221> misc_feature
<222> 217..277
<223> homology
      id :AA161280
      est
<220>
<221> misc_feature
<222> 113..160
<223> homology
      id :AA161280
      est
```

179

<220>
<221> misc_feature
<222> 173..217
<223> homology
     id :AA161280
     est
<400> 52

```
aacaacttcc ggccccactg agcggtgtcc tgagccgatt acagctaggt agtggagcgc      60
cgctgcttac ctgggtgcag gagacagccg gagtcgctgg gggagctccg cgccgccgga     120
cgcccgtgac c atg tgg agg ctg ctg gct cgc gct agt gcg ccg ctc ctg      170
            Met Trp Arg Leu Leu Ala Arg Ala Ser Ala Pro Leu Leu
              -25                         -20
cgg gtg ccc ttg tca gat tcc tgg gca ctc ctc ccc gcc agt gct ggc       218
Arg Val Pro Leu Ser Asp Ser Trp Ala Leu Leu Pro Ala Ser Ala Gly
-15                 -10                 -5                  1
gta aag aca ctg ctc cca gta cca agt ttt gaa gat gtt tcc att cct       266
Val Lys Thr Leu Leu Pro Val Pro Ser Phe Glu Asp Val Ser Ile Pro
            5                   10                  15
gaa aaa ccc aag ctt aga ttt att gaa agg gca cca ctt gtg cca aaa       314
Glu Lys Pro Lys Leu Arg Phe Ile Glu Arg Ala Pro Leu Val Pro Lys
            20                  25                  30
gta aga aga gaa cct aaa aat tta agt gac ata cgg gga cct tcc act       362
Val Arg Arg Glu Pro Lys Asn Leu Ser Asp Ile Arg Gly Pro Ser Thr
    35                  40                  45
gaa gct acg gag kkk aca gaa ggc aat ttt gca atc ttg gca ttg ggt       410
Glu Ala Thr Glu Xaa Thr Glu Gly Asn Phe Ala Ile Leu Ala Leu Gly
50                  55                  60                  65
ggt ggc tac ctg cat tgg ggc cac ttt gaa atg atg cgc ctg aca atc       458
Gly Gly Tyr Leu His Trp Gly His Phe Glu Met Met Arg Leu Thr Ile
                70                  75                  80
aac cgc tct atg gac ccc aag aac atg ttt gcc ata tgg cga gta cca       506
Asn Arg Ser Met Asp Pro Lys Asn Met Phe Ala Ile Trp Arg Val Pro
            85                  90                  95
gcc cct ttc aag ccc atc act cgc aaa agt gtt ggg cat cgc atg ggg       554
Ala Pro Phe Lys Pro Ile Thr Arg Lys Ser Val Gly His Arg Met Gly
            100                 105                 110
gga ggc aaa ggt gct att gac cac tac gtg aca cct gtg aag gct ggc       602
Gly Gly Lys Gly Ala Ile Asp His Tyr Val Thr Pro Val Lys Ala Gly
    115                 120                 125
cgc mww gww gta gag atg ggt ggg cgt tgt gma ttt gaa gaa gtg caa       650
Arg Xaa Xaa Val Glu Met Gly Gly Arg Cys Xaa Phe Glu Glu Val Gln
130                 135                 140                 145
ggt ttc ctt gac cag gtt gcc cac aag ttg ccc tty gca gca aag gct       698
Gly Phe Leu Asp Gln Val Ala His Lys Leu Pro Phe Ala Ala Lys Ala
            150                 155                 160
gtg agc cgc ggg act yta gag aag atg cga aaa gat caa gag gaa aga       746
Val Ser Arg Gly Thr Leu Glu Lys Met Arg Lys Asp Gln Glu Glu Arg
            165                 170                 175
gaa mgt aac aac cag aac ccc tgg aca ttt gag cga ata gcc act gcc       794
Glu Xaa Asn Asn Gln Asn Pro Trp Thr Phe Glu Arg Ile Ala Thr Ala
            180                 185                 190
mac atg ctg ggc ata cgg aaa gta ctg agc cca tat gac ttg acc cac       842
Xaa Met Leu Gly Ile Arg Lys Val Leu Ser Pro Tyr Asp Leu Thr His
    195                 200                 205
aag ggg aaa tam tgg ggc aag tty tac atg ccc mam cgt gtg               884
Lys Gly Lys Xaa Trp Gly Lys Phe Tyr Met Pro Xaa Arg Val
210                 215                 220
tagtgagtgt aggagataac tgtatatagg stactgaaag aaggattytg catttytatt     944
```

51 /80/

```
cccctcagcc tacccactga agtytttggg tagctyttaa gccataamta aggagcagca    1004
tttgagtaga tttytgaaaa acgatgttat ttgttgattt aaaaagaaaa cwgtattttt    1064
attaaataaa atttaaacat cacttcagga aaaaaaaaaa aaa                      1107


<210> 53
<211> 500
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 199..288
<223> Von Heijne matrix
      score 5.59999990463257
      seq IVSVLALIPETTT/LT
<220>
<221> polyA_signal
<222> 464..469
<220>
<221> polyA_site
<222> 489..500
<220>
<221> misc_feature
<222> 197..412
<223> homology
       id :AA429945
      est
<220>
<221> misc_feature
<222> 61..195
<223> homology
       id :AA429945
      est
<220>
<221> misc_feature
<222> 425..488
<223> homology
       id :AA429945
      est
<220>
<221> misc_feature
<222> 197..412
<223> homology
       id :AA455042
      est
<220>
<221> misc_feature
<222> 61..195
<223> homology
       id :AA455042
      est
<220>
<221> misc_feature
<222> 425..488
<223> homology
       id :AA455042
      est
<220>
<221> misc_feature
```

```
<222> 207..412
<223> homology
      id :W93646
      est
<220>
<221> misc_feature
<222> 58..195
<223> homology
      id :W93646
      est
<220>
<221> misc_feature
<222> 425..488
<223> homology
      id :W93646
      est
<220>
<221> misc_feature
<222> 197..412
<223> homology
      id :AA516431
      est
<220>
<221> misc_feature
<222> 90..195
<223> homology
      id :AA516431
      est
<220>
<221> misc_feature
<222> 425..488
<223> homology
      id :AA516431
      est
<220>
<221> misc_feature
<222> 52..195
<223> homology
      id :W38899
      est
<220>
<221> misc_feature
<222> 197..324
<223> homology
      id :W38899
      est
<220>
<221> misc_feature
<222> 443..477
<223> homology
      id :W38899
      est
<220>
<221> misc_feature
<222> 197..338
<223> homology
      id :W52820
      est
```

53

```
<220>
<221> misc_feature
<222> 71..195
<223> homology
       id :W52820
       est
<220>
<221> misc_feature
<222> 339..401
<223> homology
       id :W52820
       est
<220>
<221> misc_feature
<222> 425..469
<223> homology
       id :W52820
       est
<220>
<221> misc_feature
<222> 40..195
<223> homology
       id :W19506
       est
<220>
<221> misc_feature
<222> 9..10,12
<223> n=a, g, c or t
<400> 53
agagctgtnn cnsaagtagg ggagggcggt gctccgcmgm ggtggcggdh tgctatcgct      60
tcgcagaacc tactcaggca gccagctgag aagagttgag ggaaagtgct gctgctgggt     120
ctgcagacgc gatggataac gtgcagccga aaataaaaca tcgccccttc tgcttcagtg     180
tgaaaggcca cgtgayag atg ctg cgg ctg gat att atc aac tca ctg gta      231
                    Met Leu Arg Leu Asp Ile Ile Asn Ser Leu Val
                    -30              -25                  -20
aca aca gta ttc atg ctc atc gta tct gtg ttg gca ctg ata cca gaa      279
Thr Thr Val Phe Met Leu Ile Val Ser Val Leu Ala Leu Ile Pro Glu
            -15              -10                  -5
acc aca aca ttg aca gtt ggt gga ggg gtg ttt gca ctt gtg aca gca      327
Thr Thr Thr Leu Thr Val Gly Gly Gly Val Phe Ala Leu Val Thr Ala
        1               5                   10
gta tgc tgt ctt gcc gac ggg gcc ctt att tac cgg aag ctt ctg ttc      375
Val Cys Cys Leu Ala Asp Gly Ala Leu Ile Tyr Arg Lys Leu Leu Phe
    15              20                  25
aat ccc agc ggt cct tac cag aaa aag cct gtg cat gaa aaa aaa gaa      423
Asn Pro Ser Gly Pro Tyr Gln Lys Lys Pro Val His Glu Lys Lys Glu
30              35                  40                  45
gtt ttg taattttata ttacttttta gtttgatact aagtattaaa catatttctg      479
Val Leu
tattcttcca aaaaaaaaaa a                                             500

<210> 54
<211> 765
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 293..385
```

54

```
aaaccttgtt gctagggacc gggcggtttg cggcaaccgt gggcactgct gaatttgaat      60
tgaggggcga gggaaaagtt ttcctcaggt gtggtgggga gagggaggcg gatgccggng     120
aaaccgtagg kacgcggtca gaaaggcgac gggctgtcgg agttggaaag ggacgcctgg     180
tttcccccca agcgaaccgg gatgggaagt gacttcaatg agattgaact tcagctggat     240
tgaaagagag gctagaagtt ccgcttgcca gcagcctcct tagtagagcg ga atg agt     298
                                                          Met Ser
                                                              -30

aat acc cac acg gtg ctt gtc tca ctt ccc cat ccg cac ccg gcc ctc      346
Asn Thr His Thr Val Leu Val Ser Leu Pro His Pro His Pro Ala Leu
                -25             -20             -15

acc tgc tgt cac ctc ggc ctc cca cac ccg gtc cgc gct ccc cgc cct      394
Thr Cys Cys His Leu Gly Leu Pro His Pro Val Arg Ala Pro Arg Pro
            -10             -5              1

ctt cct cgc gta gaa ccg tgg gat cct agg tgg cag gac tca gag cta      442
Leu Pro Arg Val Glu Pro Trp Asp Pro Arg Trp Gln Asp Ser Glu Leu
    5               10              15

agg tat cca cag gcc atg aat tcc ttc cta aat gag cgg tca tcg ccg      490
Arg Tyr Pro Gln Ala Met Asn Ser Phe Leu Asn Glu Arg Ser Ser Pro
20              25              30              35

tgc agg acc tta agg caa gaa gca tcg gct gac aga tgt gat ctc          535
Cys Arg Thr Leu Arg Gln Glu Ala Ser Ala Asp Arg Cys Asp Leu
            40              45              50

tgaacctgat agattgctga ttttatctta ttttatcctt gacttggtac aagttttggg     595
atttctgaaa agaccataca gataaccaca aatatcaaga aagtcgtctt cagtattaag     655
tagaatttag atttaggttt ccttcctgct tcccacctcc ttcgaataag gaaacgtctt     715
tgggaccaac tttatggaat aaataagctg agctgcaaaa waaaaaaaaa                 765
```

184

<221> polyA_signal
<222> 546..551
<220>
<221> polyA_site
<222> 572..584
<400> 55

```
aagacgcgcc ggtttctgcg acgcagttag cgcagtctgc tttggtgaat acacgatttg      60
gtgcagccgg ggtttggtac cgagcggaga ggagatgcac acggcactcg agtgtgagga     120
aaaatagaa atg aag gta cat atg cac aca aaa ttt tgc ctc att tgt ttg     171
          Met Lys Val His Met His Thr Lys Phe Cys Leu Ile Cys Leu
              -20               -15                   -10
ctg aca ttt att ttt cat cat tgc aac cat tgc cat gaa gaa cat gac      219
Leu Thr Phe Ile Phe His His Cys Asn His Cys His Glu Glu His Asp
     -5                1               5                   10
cat ggc cct gaa gcg ctt cac aga cag cat cgt gga atg aca gaa ttg      267
His Gly Pro Glu Ala Leu His Arg Gln His Arg Gly Met Thr Glu Leu
             15                20                  25
gag cca agc aaa ttt tca aag caa gct gct gaa aat gaa aaa aaa tac      315
Glu Pro Ser Lys Phe Ser Lys Gln Ala Ala Glu Asn Glu Lys Lys Tyr
             30                35                  40
tat att gaa aaa ctt ttt gag cgt tat ggt gaa aat gga aga tta tcc      363
Tyr Ile Glu Lys Leu Phe Glu Arg Tyr Gly Glu Asn Gly Arg Leu Ser
             45                50                  55
ttt ttt ggt ttg gag aaa ctt tta aca aac ttg ggc ctt gga gag aga      411
Phe Phe Gly Leu Glu Lys Leu Leu Thr Asn Leu Gly Leu Gly Glu Arg
   60                65                  70
aaa gta gtt gag att aat cat gag gat ctt ggc cac gat cat gtt tct      459
Lys Val Val Glu Ile Asn His Glu Asp Leu Gly His Asp His Val Ser
75                80                  85                  90
cat tta agg tat ttt ggc agt tca aga ggg aaa gca ttt tca ctc aca      507
His Leu Arg Tyr Phe Gly Ser Ser Arg Gly Lys Ala Phe Ser Leu Thr
             95                100                 105
taaccaccca gcattcccat aatcatttaa attcagaaaa tcaaaactgt gaccagtgta     567
wtccacaaaa aaaaaaa                                                    584
```

<210> 56
<211> 1387
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 191..325
<223> Von Heijne matrix
      score 4.59999990463257
      seq VLVYLVTAERVWS/DD
<220>
<221> polyA_signal
<222> 1348..1353
<220>
<221> polyA_site
<222> 1374..1387
<220>
<221> misc_feature
<222> 1258..1372
<223> homology
      id :AA417826
      est
<220>

```
<221> misc_feature
<222> 791..887
<223> homology
      id :AA417826
      est
<220>
<221> misc_feature
<222> 94..524
<223> homology
      id :AA235826
      est
<220>
<221> misc_feature
<222> 44..94
<223> homology
      id :AA235826
      est
<220>
<221> misc_feature
<222> 1258..1372
<223> homology
      id :AA236941
      est
<220>
<221> misc_feature
<222> 935..1279
<223> homology
      id :AA480326
      est
<220>
<221> misc_feature
<222> 1258..1372
<223> homology
      id :AA480326
      est
<220>
<221> misc_feature
<222> 724..1148
<223> homology
      id :AA234245
      est
<220>
<221> misc_feature
<222> 944..1279
<223> homology
      id :AA479344
      est
<220>
<221> misc_feature
<222> 1258..1372
<223> homology
      id :AA479344
      est
<220>
<221> misc_feature
<222> 1070..1212
<223> homology
      id :AA133636
```

```
          est
<220>
<221> misc_feature
<222> 1258..1372
<223> homology
        id :AA133636
        est
<220>
<221> misc_feature
<222> 938..1054
<223> homology
        id :AA133636
        est
<220>
<221> misc_feature
<222> 94..436
<223> homology
        id :AA133635
        est
<220>
<221> misc_feature
<222> 32..94
<223> homology
        id :AA133635
        est
<220>
<221> misc_feature
<222> 895..1273
<223> homology
        id :AA479453
        est
<220>
<221> misc_feature
<222> 1258..1371
<223> homology
        id :AA253214
        est
<220>
<221> misc_feature
<222> 94..268
<223> homology
        id :AA482378
        est
<220>
<221> misc_feature
<222> 946
<223> n=a, g, c or t
<400> 56
actcccaggc tgggccagca cacccggcag gctctgtcct ggaaacaggc ttcaacgggc      60
ttccccgaaa accttccccg cttctggata tgaavattca agctgcttgc tgagtcctat     120
tgccggctgc tgggagccag gagagccctg aggagtagtc actcagtagc agctgacgcg     180
tgggtccacc atg aac tgg agt atc ttt gag gga ctc ctg agt ggg gtc        229
            Met Asn Trp Ser Ile Phe Glu Gly Leu Leu Ser Gly Val
            -45                 -40                 -35
aac aag tac tcc aca gcc ttt ggg cgc atc tgg ctg tct ctg gtc ttc       277
Asn Lys Tyr Ser Thr Ala Phe Gly Arg Ile Trp Leu Ser Leu Val Phe
        -30                 -25                 -20
atc ttc cgc gtg ctg gtg tac ctg gtg acg gcc gag cgt gtg tgg agt       325
```

58 187

```
Ile Phe Arg Val Leu Val Tyr Leu Val Thr Ala Glu Arg Val Trp Ser
    -15                     -10                 -5
gat gac cac aag gac ttc gac tgc aat act cgc cag ccc ggc tgc tcc      373
Asp Asp His Lys Asp Phe Asp Cys Asn Thr Arg Gln Pro Gly Cys Ser
1               5                   10                  15
aac gtc tgc ttt gat gag ttc ttc cct gtg tcc cat gtg cgc ctc tgg      421
Asn Val Cys Phe Asp Glu Phe Phe Pro Val Ser His Val Arg Leu Trp
            20                  25                  30
gcc ctg cag ctt atc ctg gtg aca tgc ccc tca ctg ctc gtg gtc atg      469
Ala Leu Gln Leu Ile Leu Val Thr Cys Pro Ser Leu Leu Val Val Met
        35                  40                  45
cac gtg gcc tac cgg gag gtt cag gag aag agg cac cga gaa gcc cat      517
His Val Ala Tyr Arg Glu Val Gln Glu Lys Arg His Arg Glu Ala His
    50                  55                  60
ggg gag aac agt ggg cgc ctc tac ctg aac ccc ggc aag aar cgg ggt      565
Gly Glu Asn Ser Gly Arg Leu Tyr Leu Asn Pro Gly Lys Lys Arg Gly
65                  70                  75                  80
ggg ctc tgg tgg aca tat gtc tgc agc cta gtg ttc aag gcg agc gtg      613
Gly Leu Trp Trp Thr Tyr Val Cys Ser Leu Val Phe Lys Ala Ser Val
                85                  90                  95
gac atc gcc ttt ctc tat gtg ttc cac tca ttc tac ccc aaa tat atc      661
Asp Ile Ala Phe Leu Tyr Val Phe His Ser Phe Tyr Pro Lys Tyr Ile
            100                 105                 110
ctc cct cct gtg gtc aag tgc cac gca gat cca tgt ccc aat ata gtg      709
Leu Pro Pro Val Val Lys Cys His Ala Asp Pro Cys Pro Asn Ile Val
        115                 120                 125
gac tgc ttc atc tcc aag ccc tca gag aag aac att ttc acc ctc ttc      757
Asp Cys Phe Ile Ser Lys Pro Ser Glu Lys Asn Ile Phe Thr Leu Phe
    130                 135                 140
atg gtg gcc aca gct gcc atc tgc atc ctg ctc aac ctc gtg gag ctc      805
Met Val Ala Thr Ala Ala Ile Cys Ile Leu Leu Asn Leu Val Glu Leu
145                 150                 155                 160
atc tac ctg gtg agc aag aga tgc cac gag tgc ctg gca gca agg aaa      853
Ile Tyr Leu Val Ser Lys Arg Cys His Glu Cys Leu Ala Ala Arg Lys
                165                 170                 175
gct caa gcc atg kgc aca ggt cat cac ccc cav gat acc acy ttt tcc      901
Ala Gln Ala Met Xaa Thr Gly His His Pro Xaa Asp Thr Thr Phe Ser
            180                 185                 190
kgc aaa caa gas gac ytc ytt tcg ggk gac ytc atc ttt ctg ggn tca      949
Xaa Lys Gln Xaa Asp Xaa Xaa Ser Gly Asp Xaa Ile Phe Leu Gly Ser
        195                 200                 205
gac agt cat cyt cct ytc tta cca gac cgc ccc cga gac cat gtg aag      997
Asp Ser His Xaa Pro Xaa Leu Pro Asp Arg Pro Arg Asp His Val Lys
    210                 215                 220
aaa acc aty ttg tgaggggctg cctggamtgg tytggcaggt tgggcctgga         1049
Lys Thr Ile Leu
225
tggggaggct ytagcatyty tcataggtgc aacctgagag tggggagct aagccatgag    1109
gtaggggcag gcaagagaga ggattcagac gytytgggag ccagttccta gtcctcaamt   1169
ccagccacct gccccagsth dacggcamtg ggccagttcc ccctytgsty tgcagstcgg   1229
tttccttty tagaatggaa atagtgaggg ccaatgccca gggttggagg gaggagggcg    1289
ttcatagaag aacacacatg cgggcacctt catygtgtgt ggcccactgt cagaacttaa   1349
taaaagtcaa mtcatttgct ggttaaaaaa aaaaaaaa                           1387
```

<210> 57
<211> 1385
<212> DNA
<213> Homo sapiens

59

<220>
<221> sig_peptide
<222> 141..251
<223> Von Heijne matrix
       score 4
       seq PLSLDCGHSLCRA/CI
<220>
<221> polyA_signal
<222> 1354..1359
<220>
<221> polyA_site
<222> 1375..1385
<220>
<221> misc_feature
<222> 1183..1240
<223> homology
       id :AA463623
       est
<220>
<221> misc_feature
<222> 176..239
<223> homology
       id :AA258927
       est
<220>
<221> misc_feature
<222> 803..854
<223> homology
       id :AA286417
       est
<220>
<221> misc_feature
<222> 1183..1213
<223> homology
       id :AA608077
       est
<400> 57
aacacccacc ctggctttc ttcacctctt caaccaggag ccgagatttc tgttgctctg   60
aagccatcca ggggtctta accagaagag agaggagagc ctcaggagtt aggaccagaa   120
gaagccaggg aagcagtgca atg gct tca aaa atc ttg ctt aac gta caa gag   173
                        Met Ala Ser Lys Ile Leu Leu Asn Val Gln Glu
                                    -35   -30
gag gtg acc tgt ccc atc tgc ctg gag ctg ttg aca gaa ccc ttg agt   221
Glu Val Thr Cys Pro Ile Cys Leu Glu Leu Leu Thr Glu Pro Leu Ser
    -25   -20   -15
cta gac tgt ggc cac agc ctc tgc cga gcc tgc atc act gtg agc aac   269
Leu Asp Cys Gly His Ser Leu Cys Arg Ala Cys Ile Thr Val Ser Asn
-10   -5   1   5
aag gag gca gtg acc agc atg gga gga aaa agc agc tgt cct gtg tgt   317
Lys Glu Ala Val Thr Ser Met Gly Gly Lys Ser Ser Cys Pro Val Cys
            10   15   20
ggt atc agt tac tca ttt gaa cat cta cag gct aat cag cat cgg gcc   365
Gly Ile Ser Tyr Ser Phe Glu His Leu Gln Ala Asn Gln His Arg Ala
        25   30   35
aac ata gtg gag aga ctc aag gag gtc aag ttg agc cca gac aat ggg   413
Asn Ile Val Glu Arg Leu Lys Glu Val Lys Leu Ser Pro Asp Asn Gly
    40   45   50
aag aag aga gat ctc tgt gat cat cat gga gag aaa ctc cta ctc ttc   461

60

```
Lys Lys Arg Asp Leu Cys Asp His His Gly Glu Lys Leu Leu Leu Phe
55                      60                  65                  70
tgt aag gag gat agg aaa gtc att tgc tgg ctt tgt gag cgg tct cag       509
Cys Lys Glu Asp Arg Lys Val Ile Cys Trp Leu Cys Glu Arg Ser Gln
                75                  80                  85
gag cac cgt ggt cac cac aca ggt cct cac gga gga agt att caa gga       557
Glu His Arg Gly His His Thr Gly Pro His Gly Gly Ser Ile Gln Gly
                90                  95                  100
atg tca gga gaa act cca ggc agt cct caa gag gct gaa gaa gga aga       605
Met Ser Gly Glu Thr Pro Gly Ser Pro Gln Glu Ala Glu Glu Gly Arg
            105                 110                 115
gga gga agc tgagaagctg gaagctgaca tcagagaaga gaaaacttcc              654
Gly Gly Ser
            120
tggaagtatc aggtacaaac tgagagacaa aggatacaaa cagaatttga tcagcttaga    714
agcatcctaa ataatgagga gcagagagag ctgcaaagat tggaagaaga agaaaagaag    774
acgctggata agtttgcaga ggctgaggat gagctagttc agcagaagca gttggtgaga    834
gagctcatct cagatgtgga gtgtcggagt cagtggtcaa caatggagct gctgcaggac    894
atgagtggaa tcatgaaatg gagtgagatc tggaggctga aaaagccaaa aatggtttcc    954
aagaaactga agactgtatt ccatgctcca gatctgagta ggatgctgcr aatgtttaga   1014
ggaactgaca gctgtccggt gctactgggt ggatgtcaca ctgaattcag tcaacctaaa   1074
tttgaatckt gtcctttcag aagatcagag acaagtgata tctgtgccaa tttggccttt   1134
tcagtgttat aattatggtg tkbttgggat cccaatattt btcctsstgg gaaacattac   1194
tgggaagtgg acgtgtccaa gaaaactgcc tggatcctgg gggtatactg tagaacatat   1254
tcccgccata tgaagtatgt tgttagaaga tgtgcaaaty gtcaaaatbt ttacaccaaa   1314
tacagacctc tatttggsta ctgggttata gggttacaga ataaatgtaa gtatggtgcc   1374
aaaaaaaaaa a                                                        1385

<210> 58
<211> 1497
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 212..268
<223> Von Heijne matrix
      score 8.60000038146973
      seq LLWLALACSPVHT/TL
<220>
<221> polyA_signal
<222> 1465..1470
<220>
<221> polyA_site
<222> 1489..1497
<220>
<221> misc_feature
<222> 958..1110
<223> homology
      id :W72124
      est
<220>
<221> misc_feature
<222> 1362..1488
<223> homology
      id :W72124
      est
<220>
<221> misc_feature
```

```
<222> 1202..1312
<223> homology
      id :W72124
      est
<220>
<221> misc_feature
<222> 1115..1190
<223> homology
      id :W72124
      est
<220>
<221> misc_feature
<222> 1312..1370
<223> homology
      id :W72124
      est
<220>
<221> misc_feature
<222> 653..942
<223> homology
      id :AA009415
      est
<220>
<221> misc_feature
<222> 454..605
<223> homology
      id :AA009415
      est
<220>
<221> misc_feature
<222> 598..639
<223> homology
      id :AA009415
      est
<220>
<221> misc_feature
<222> 805..1032
<223> homology
      id :AA088502
      est
<220>
<221> misc_feature
<222> 633..807
<223> homology
      id :AA088502
      est
<220>
<221> misc_feature
<222> 598..639
<223> homology
      id :AA088502
      est
<220>
<221> misc_feature
<222> 564..605
<223> homology
      id :AA088502
      est
```

```
<220>
<221> misc_feature
<222> 653..807
<223> homology
      id :AA181148
      est
<220>
<221> misc_feature
<222> 907..1046
<223> homology
      id :AA181148
      est
<220>
<221> misc_feature
<222> 475..605
<223> homology
      id :AA181148
      est
<220>
<221> misc_feature
<222> 598..639
<223> homology
      id :AA181148
      est
<220>
<221> misc_feature
<222> 1069..1190
<223> homology
      id :AA181149
      est
<220>
<221> misc_feature
<222> 1362..1475
<223> homology
      id :AA181149
      est
<220>
<221> misc_feature
<222> 1202..1312
<223> homology
      id :AA181149
      est
<220>
<221> misc_feature
<222> 1312..1370
<223> homology
      id :AA181149
      est
<220>
<221> misc_feature
<222> 72,93
<223> n=a, g, c or t
<400> 58
atccggcgcg ctggagcgtt ttccggccgt gcgtttgtgg ccgtccggcc tccctgacat       60
gcagatttcc anssagaaga cagagaagga gcnagtggtc atggaatggg ctggggtcaa      120
agactgggtg cctgggagct gaggcagcca ccgtttcagc ctggccagcc ctctggaccc      180
cgaggttgga ccctactgtg acacacctac c atg cgg aca ctc ttc aac ctc        232
                                    Met Arg Thr Leu Phe Asn Leu
```

```
ctc tgg ctt gcc ctg gcc tgc agc cct gtt cac act acc ctg tca aag      280
Leu Trp Leu Ala Leu Ala Cys Ser Pro Val His Thr Thr Leu Ser Lys
        -10                 -5                  1
tca gat gcc asa aaa ccg cct caa aga cgc tgc tgg aga aga gtc agt      328
Ser Asp Ala Xaa Lys Pro Pro Gln Arg Arg Cys Trp Arg Arg Val Ser
5                   10                  15                  20
ttt cag ata agc cgg tgc aar acc ggg gtt tgg tgg tgacggacct          374
Phe Gln Ile Ser Arg Cys Lys Thr Gly Val Trp Trp
                25                  30
caaagctgag agtgtggttc ttgagcatcg cagctactgc tcggcaaagg cccgggacag    434
acactttgct ggggatgtac tgggctatgt cactccatgg aacagccatg gctacgatgt    494
caccaaggtc tttgggagca agttcacaca gatctcaccc gtctggctgc agttgaagag    554
acgtggccgt gagatgtttg aggtcacggg cctccacgac gtggaccaag ggtggatgcg    614
agctgtcagg aagcatgcca agggcctgca catagtgcct cggctcctgt ttgaggactg    674
gacttacgat gatttccgga acgtcttaga cagtgaggat gagatagagg agctgagcaa    734
gaccgtggtc caggtggcaa agaaccagca tttcgatggc ttcgtggtgg aggtctggaa    794
ccagctgcta agccagaagc gcgtgggcct catccacatg ctcacccact ggccgaggc    854
cctgcaccag gcccggctgc tggccctcct ggtcatcccg cctgccatca ccccegggac    914
cgaccagctg ggcatgttca cgcacaagga gtttgagcag ctggcccccg tgctggatgg    974
tttcagcctc atgacctacg actactctac agcgcatcag cctggcccta atgcacccct    1034
gtcctgggtt cgagcctgcg tccaggtcct ggacccggaa gtccaagtgg cgaagcaaaa    1094
tcctcctggg gctcaacttc tatggtatgg actacgcgac ctccaaggat gcccgtgagc    1154
ctgttgtcgg ggccaggtac atccagacac tgaadggacc acaggccccg ggaatggtgt    1214
gggacagcca ggcctcagag cacttcttcg agtacaagaa gagccgcagt gggaggcacg    1274
tcgtcttcta cccaaccctg aagtccctgc aggtgcgggc tggagctggc ccgggagctg    1334
ggcgttgggg tctctatytg ggagctgggc cagggcctgg actacttyta cgacctgcty    1394
taggtgggca ttgcggcctc cgcggtggac gtgttytttt ytaagccatg gagtgagtga    1454
gcaggtgtga aatacaggcc tccactccgt ttgcaaaaaa aaa                      1497
```

```
<210> 59
<211> 1570
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 147..248
<223> Von Heijne matrix
      score 4.30000019073486
      seq QLFAFLNLLPVEA/DI
<220>
<221> polyA_signal
<222> 1538..1543
<220>
<221> polyA_site
<222> 1558..1570
<220>
<221> misc_feature
<222> 466..968
<223> homology
      id :AA506103
      est
<220>
<221> misc_feature
<222> 142..664
<223> homology
      id :AA237105
      est
```

```
<220>
<221> misc_feature
<222> 114..269
<223> homology
      id :AA317201
      est
<220>
<221> misc_feature
<222> 2..122
<223> homology
      id :AA317201
      est
<220>
<221> misc_feature
<222> 401..443
<223> homology
      id :AA317201
      est
<220>
<221> misc_feature
<222> 103..385
<223> homology
      id :T80259
      est
<220>
<221> misc_feature
<222> 21..120
<223> homology
      id :T80259
      est
<220>
<221> misc_feature
<222> 109..459
<223> homology
      id :N32697
      est
<220>
<221> misc_feature
<222> 45..87
<223> homology
      id :N32697
      est
<220>
<221> misc_feature
<222> 92..122
<223> homology
      id :N32697
      est
<220>
<221> misc_feature
<222> 1220..1409
<223> homology
      id :AA449621
      est
<220>
<221> misc_feature
<222> 928..1092
<223> homology
```

194

```
            id :AA449621
        est
<220>
<221> misc_feature
<222> 1178..1222
<223> homology
            id :AA449621
        est
<220>
<221> misc_feature
<222> 1220..1545
<223> homology
            id :N34685
        est
<220>
<221> misc_feature
<222> 1168..1222
<223> homology
            id :N34685
        est
<220>
<221> misc_feature
<222> 1220..1545
<223> homology
            id :N22990
        est
<220>
<221> misc_feature
<222> 1178..1222
<223> homology
            id :N22990
        est
<220>
<221> misc_feature
<222> 114..325
<223> homology
            id :AA330462
        est
<220>
<221> misc_feature
<222> 18..122
<223> homology
            id :AA330462
        est
<220>
<221> misc_feature
<222> 135..475
<223> homology
            id :HUMEST5H12
        est
<220>
<221> misc_feature
<222> 1093,1128,1135,1147
<223> n=a, g, c or t
<400> 59
agtcgtccct gctagtactc cgggctgtgg gggtcggtgc ggatattcag tcatgaaatc      60
agggtaggga cttctcccgc agcgacgcgg ctggcaagac tgtttgtgtt gcgggggccg     120
gaacttcaag gtgattttac aacgag atg ctg ctc tcc ata ggg atg ctc atg     173
```

66

```
                              Met Leu Leu Ser Ile Gly Met Leu Met
                                                   -30
ctg tca gcc aca caa gtc tac acc atc ttg act gtc cag ctc ttt gca        221
Leu Ser Ala Thr Gln Val Tyr Thr Ile Leu Thr Val Gln Leu Phe Ala
-25              -20                 -15                     -10
ttc tta aac cta ctg cct gta gaa gca gac att tta gca tat aac ttt        269
Phe Leu Asn Leu Leu Pro Val Glu Ala Asp Ile Leu Ala Tyr Asn Phe
              -5                  1                   5
gaa aat gca tct cag aca ttt gat gac ctc ccc gca ara ttt ggt tat        317
Glu Asn Ala Ser Gln Thr Phe Asp Asp Leu Pro Ala Xaa Phe Gly Tyr
              10                  15                  20
aga ctt cca gct gaa ggt tta aag ggt ttt tta att aac tca aaa cca        365
Arg Leu Pro Ala Glu Gly Leu Lys Gly Phe Leu Ile Asn Ser Lys Pro
    25                  30                  35
gag aat gcc tgt gaa ccc ata gtg cct cca cca gta aaa gac aat tca        413
Glu Asn Ala Cys Glu Pro Ile Val Pro Pro Pro Val Lys Asp Asn Ser
40                  45                  50                      55
tct ggc act ttc atc gtg tta att ara ara ctt gat tgt aat ttt gat        461
Ser Gly Thr Phe Ile Val Leu Ile Xaa Xaa Leu Asp Cys Asn Phe Asp
              60                  65                  70
ata aag gtt tta aat gca cag aga gca gga tac aag gca gcc ata gtt        509
Ile Lys Val Leu Asn Ala Gln Arg Ala Gly Tyr Lys Ala Ala Ile Val
              75                  80                  85
cac aat gtt gat tct gat gac ctc att agc atg gga tcc aac gac att        557
His Asn Val Asp Ser Asp Asp Leu Ile Ser Met Gly Ser Asn Asp Ile
              90                  95                  100
gag gta cta aag aaa att gac att cca tct gtc ttt att ggt gaa tca        605
Glu Val Leu Lys Lys Ile Asp Ile Pro Ser Val Phe Ile Gly Glu Ser
    105                 110                 115
tca gct agt tct ctg aaa gat gaa ttc aca tak gaa aaa ggg ggc cac        653
Ser Ala Ser Ser Leu Lys Asp Glu Phe Thr Xaa Glu Lys Gly Gly His
120                 125                 130                     135
ctt atc tta gtt cca gaa ttt agt ctt cct ttg gaa tac tac cta att        701
Leu Ile Leu Val Pro Glu Phe Ser Leu Pro Leu Glu Tyr Tyr Leu Ile
              140                 145                 150
ccc ttc ctt atc atr gtg ggc atc tgt ctc atc ttg ata gtc att ttc        749
Pro Phe Leu Ile Xaa Val Gly Ile Cys Leu Ile Leu Ile Val Ile Phe
              155                 160                 165
atg atc aca aaa ttg tcc agg gat aga cat aga gct aga aga aac aga        797
Met Ile Thr Lys Leu Ser Arg Asp Arg His Arg Ala Arg Arg Asn Arg
              170                 175                 180
ctt cgt aaa gat caa ctt aag aaa ctt cct gta cat aaa ttc aag aaa        845
Leu Arg Lys Asp Gln Leu Lys Lys Leu Pro Val His Lys Phe Lys Lys
    185                 190                 195
gga gat gag tat gat gta tgt gcc att tgt ttg gat gag tat gaa gat        893
Gly Asp Glu Tyr Asp Val Cys Ala Ile Cys Leu Asp Glu Tyr Glu Asp
200                 205                 210                     215
gga gac aaa ctc aga atc ctt ccc tgt tcc cat gct tat cat tgc aag        941
Gly Asp Lys Leu Arg Ile Leu Pro Cys Ser His Ala Tyr His Cys Lys
              220                 225                 230
tgt gta gac cct tgg cta act aaa acc aaa aaa acc tgt cca gtg tgc        989
Cys Val Asp Pro Trp Leu Thr Lys Thr Lys Lys Thr Cys Pro Val Cys
              235                 240                 245
agg caa aaa gtt gtt cct tct caa ggc gat tca gac tct gac aca gac        1037
Arg Gln Lys Val Val Pro Ser Gln Gly Asp Ser Asp Ser Asp Thr Asp
    250                 255                 260
agt agt caa gaa gaa aat gaa gtg aca gaa cat acc cct tta ctg aga        1085
Ser Ser Gln Glu Glu Asn Glu Val Thr Glu His Thr Pro Leu Leu Arg
```

67

```
         265                      270                      275
cct tta gnc ttc tgt cag tgc cca rgt cam ttt ggg gct tta ntc gga      1133
Pro Leu Xaa Phe Cys Gln Cys Pro Xaa Xaa Phe Gly Ala Leu Xaa Gly
280                      285                      290                      295
ant ccc gct cac ant cag aak cat gac aga atc att cag act ast gag      1181
Xaa Pro Ala His Xaa Gln Xaa His Asp Arg Ile Ile Gln Thr Xaa Glu
                         300                      305                      310
gaa gac gac aat gaa gat act gac agt agt gat gca gaa gaa              1223
Glu Asp Asp Asn Glu Asp Thr Asp Ser Ser Asp Ala Glu Glu
                    315                      320                      325
tgaaattaat gaacatgatg tcgtggtcca gttgcagcct aatggtgaac gggattacaa   1283
catagcaaat actgtttgac tttcagaaga tgattggttt atttcccttt aaaatgatta   1343
ggtatatact gtaatttgat tttttgctcc cttaaaagat ttytgtagaa ataacttatt   1403
ttttagtact ytacagttta atcaaattac tgaaacagga cttttgatyt ggtatttatc   1463
tgccaagaat atacttcatt cactaataat agactggtgc tgtaactcaa gcatcaattc   1523
agctyttytt ttggaatgaa agtatagcca aaacaaaaaa aaaaaaa                  1570


<210> 60
<211> 1022
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 112..237
<223> Von Heijne matrix
      score 7.19999980926514
      seq ILFSLSFLLVIIT/FP
<220>
<221> polyA_signal
<222> 976..981
<220>
<221> polyA_site
<222> 1010..1022
<400> 60
aatactttct cctctcccct ctcccaagca catctgagtt gctgcctgtt cttcacactt     60
agctccaaac ccatgaaaaa ttgccaagta taaaagcttc tcaagaatga g atg gat    117
                                                          Met Asp
tct agg gtg tct tca cct gag aag caa gat aaa gag aat ttc gtg ggt     165
Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys Glu Asn Phe Val Gly
-40                      -35                      -30                      -25
gtc aac aat aaa cgg ctt ggt gta tgt ggc tgg atc ctg ttt tcc ctc     213
Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp Ile Leu Phe Ser Leu
                    -20                      -15                      -10
tct ttc ctg ttg gtg atc att acc ttc ccc atc tcc ata tgg atg tgc     261
Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile Ser Ile Trp Met Cys
               -5                       1                        5
ttg aag atc att aag gag tat gaa cgt gct gtt gta ttc cgt ctg gga     309
Leu Lys Ile Ile Lys Glu Tyr Glu Arg Ala Val Val Phe Arg Leu Gly
          10                       15                       20
cgc atc caa gct gac aaa gcc aag ggg cca ggt ttg atc ctg gtc ctg     357
Arg Ile Gln Ala Asp Lys Ala Lys Gly Pro Gly Leu Ile Leu Val Leu
25                       30                       35                       40
cca tgc ata gat gtg ttt gtc aag gtt gac ctc cga aca gtt act tgc     405
Pro Cys Ile Asp Val Phe Val Lys Val Asp Leu Arg Thr Val Thr Cys
                    45                       50                       55
aac att cct cca caa gag atc ctc acc aga gac tcc gta act act cag     453
Asn Ile Pro Pro Gln Glu Ile Leu Thr Arg Asp Ser Val Thr Thr Gln
                    60                       65                       70
```

68

```
gta gat gga gtt gtc tat tac aga atc tat agt gct gtc tca gca gtg        501
Val Asp Gly Val Val Tyr Tyr Arg Ile Tyr Ser Ala Val Ser Ala Val
        75                  80                  85
gct aat gtc aac gat gtc cat caa gca aca ttt ctg ctg gct caa acc        549
Ala Asn Val Asn Asp Val His Gln Ala Thr Phe Leu Leu Ala Gln Thr
        90                  95                  100
act ctg aga aat gtc tta ggg aca cag acc ttg tcc cag atc tta gct        597
Thr Leu Arg Asn Val Leu Gly Thr Gln Thr Leu Ser Gln Ile Leu Ala
105                 110                 115                 120
gga cga gaa gag atc gcc cat agc atc cag act tta ctt gat gat gcc        645
Gly Arg Glu Glu Ile Ala His Ser Ile Gln Thr Leu Leu Asp Asp Ala
                    125                 130                 135
acc gaa ctg tgg ggg atc cgg gtg gcc cga gtg gaa atc aaa gat gtt        693
Thr Glu Leu Trp Gly Ile Arg Val Ala Arg Val Glu Ile Lys Asp Val
                140                 145                 150
cgg att ccc gtg cag ttg cag aga tcc atg gca gcc gag gct gag gcc        741
Arg Ile Pro Val Gln Leu Gln Arg Ser Met Ala Ala Glu Ala Glu Ala
            155                 160                 165
acc cgg gaa gcg aga gcc aag gtc ctt gca gct gaa gga gaa atg agt        789
Thr Arg Glu Ala Arg Ala Lys Val Leu Ala Ala Glu Gly Glu Met Ser
        170                 175                 180
gct tcc aaa tcc ctg aag tca gcc tcc atg gtg ctg gct gag tct ccc        837
Ala Ser Lys Ser Leu Lys Ser Ala Ser Met Val Leu Ala Glu Ser Pro
185                 190                 195                 200
ata gct ctc cag ctg cgc tac ctg cag acc ttg agc acg gta gcc acc        885
Ile Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu Ser Thr Val Ala Thr
                205                 210                 215
gag aag aat tct acg att gtg ttt cct ctg ccc atg aat ata cta gag        933
Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro Met Asn Ile Leu Glu
                220                 225                 230
ggc att ggt ggc gtc agc tat gat aac cac aag aag ctt cca aat aaa        981
Gly Ile Gly Gly Val Ser Tyr Asp Asn His Lys Lys Leu Pro Asn Lys
        235                 240                 245
gcc tgaggtcctc ttgcggtagt cagctaaaaa aaaaaaaa                          1022
Ala

<210> 61
<211> 615
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 239..316
<223> Von Heijne matrix
      score 3.90000009536743
      seq ITWVSLFIDCVMT/RK
<220>
<221> polyA_signal
<222> 586..591
<220>
<221> polyA_site
<222> 603..615
<220>
<221> misc_feature
<222> 341..574
<223> homology
      id :AA453275
      est
```

69

```
<220>
<221> misc_feature
<222> 174..332
<223> homology
        id :AA453275
        est
<220>
<221> misc_feature
<222> 85..171
<223> homology
        id :AA453275
        est
<220>
<221> misc_feature
<222> 341..574
<223> homology
        id :AA149631
        est
<220>
<221> misc_feature
<222> 170..339
<223> homology
        id :AA149631
        est
<220>
<221> misc_feature
<222> 43..123
<223> homology
        id :AA149631
        est
<220>
<221> misc_feature
<222> 88..339
<223> homology
        id :AA588414
        est
<220>
<221> misc_feature
<222> 341..574
<223> homology
        id :AA588414
        est
<220>
<221> misc_feature
<222> 1..345
<223> homology
        id :AA156847
        est
<220>
<221> misc_feature
<222> 342..414
<223> homology
        id :AA156847
        est
<220>
<221> misc_feature
<222> 341..574
<223> homology
```

```
         id :AA501739
         est
<220>
<221> misc_feature
<222> 110..339
<223> homology
         id :AA501739
         est
<220>
<221> misc_feature
<222> 341..574
<223> homology
         id :AA131792
         est
<220>
<221> misc_feature
<222> 153..259
<223> homology
         id :AA131792
         est
<220>
<221> misc_feature
<222> 259..339
<223> homology
         id :AA131792
         est
<220>
<221> misc_feature
<222> 59..338
<223> homology
         id :AA131842
         est
<220>
<221> misc_feature
<222> 344..415
<223> homology
         id :AA131842
         est
<220>
<221> misc_feature
<222> 400..434
<223> homology
         id :AA131842
         est
<220>
<221> misc_feature
<222> 341..574
<223> homology
         id :AA152042
         est
<220>
<221> misc_feature
<222> 183..339
<223> homology
         id :AA152042
         est
<400> 61
atctttgaag aagaagaagt tgaatttatc agtgtgcctg tcccagagtt tgcagatagt      60
```

71

```
gatcctgcca acattgttca tgactttaac aagaaactta cagcctattt agatcttaac      120
ctggataagt gctatgtgat ccctctgaac acttccattg ttatgccacc cagaaaccta      180
ctggagttac ttattaacat caaggctgga acctatttgc ctcagtccta tctgattc        238
atg agc aca tgg tta tta ctg atc gca ttg aaa aca ttg atc acc tgg        286
Met Ser Thr Trp Leu Leu Leu Ile Ala Leu Lys Thr Leu Ile Thr Trp
    -25               -20               -15
gtt tct tta ttt atc gac tgt gtc atg aca agg aaa ctt aca aac tgc        334
Val Ser Leu Phe Ile Asp Cys Val Met Thr Arg Lys Leu Thr Asn Cys
-10               -5                1                 5
aac gct aga gaa act att aaa ggt att cag aaa cgt gaa gcc agc aat        382
Asn Ala Arg Glu Thr Ile Lys Gly Ile Gln Lys Arg Glu Ala Ser Asn
          10                15                20
tgt ttc gca att cgg cat ttt gaa aac aaa ttt gcc gtg gaa act tta        430
Cys Phe Ala Ile Arg His Phe Glu Asn Lys Phe Ala Val Glu Thr Leu
          25                30                35
att tgt tct tgaacagtca agaaaaacat tattgaggaa aattaatatc                479
Ile Cys Ser
    40
acagcataac cccacccttt acattttgtg cagtgattat tttttaaagt cttctttcat      539
gtaagtagca aacagggctt tactatcttt tcatctcatt aattcaatta aaaccattac      599
cccaaaaaaa aaaaaa                                                       615
```

<210> 62
<211> 804
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 157..345
<223> Von Heijne matrix
      score 3.5
      seq GLVCAGLADMARP/AE
<220>
<221> polyA_signal
<222> 771..776
<220>
<221> polyA_site
<222> 791..804
<220>
<221> misc_feature
<222> 244..789
<223> homology
      id :AA576425
      est
<220>
<221> misc_feature
<222> 286..790
<223> homology
      id :AA236527
      est
<220>
<221> misc_feature
<222> 287..790
<223> homology
      id :AA435919
      est
<220>
<221> misc_feature

```
<222> 520..790
<223> homology
      id :AA165350
      est
<220>
<221> misc_feature
<222> 389..522
<223> homology
      id :AA165350
      est
<220>
<221> misc_feature
<222> 336..386
<223> homology
      id :AA165350
      est
<220>
<221> misc_feature
<222> 326..790
<223> homology
      id :AA490322
      est
<220>
<221> misc_feature
<222> 326..790
<223> homology
      id :AA490310
      est
<220>
<221> misc_feature
<222> 515..780
<223> homology
      id :AA164559
      est
<220>
<221> misc_feature
<222> 325..522
<223> homology
      id :AA164559
      est
<220>
<221> misc_feature
<222> 350..790
<223> homology
      id :AA427895
      est
<220>
<221> misc_feature
<222> 378..790
<223> homology
      id :AA532390
      est
<220>
<221> misc_feature
<222> 186..382
<223> homology
      id :AA082259
      est
```

73

```
<220>
<221> misc_feature
<222> 61..141
<223> homology
       id :AA082259
       est
<220>
<221> misc_feature
<222> 426..478
<223> homology
       id :AA082259
       est
<220>
<221> misc_feature
<222> 29..61
<223> homology
       id :AA082259
       est
<220>
<221> misc_feature
<222> 389..790
<223> homology
       id :AA157009
       est
<220>
<221> misc_feature
<222> 425..790
<223> homology
       id :AA034912
       est
<220>
<221> misc_feature
<222> 186..430
<223> homology
       id :AA428006
       est
<220>
<221> misc_feature
<222> 59..132
<223> homology
       id :AA428006
       est
<400> 62
aacagcgggc agggaaagcc gcgggaaggg tactccaggc gagaggcgga cgcgagtcgt        60
cgtggcagga aaagtgacta gctccccttc gttgtcagcc agggacgaga acacagccac       120
gctcccaccc ggctgcchaa ggatccctcg gcggcg atg tcg gcc gcc ggt gcc       174
                                          Met Ser Ala Ala Gly Ala
                                                            -60
cga ggc ctg cgg gcc acc tac cac cgg ctc ctc gat aaa gtg gag ctg       222
Arg Gly Leu Arg Ala Thr Tyr His Arg Leu Leu Asp Lys Val Glu Leu
            -55                 -50                 -45
atg ctg ccc gag aaa ttg agg ccg ttg tac aac cat cca gca ggt ccc       270
Met Leu Pro Glu Lys Leu Arg Pro Leu Tyr Asn His Pro Ala Gly Pro
    -40                 -35                 -30
aga aca gtt ttc ttc tgg gct cca att atg aaa tgg ggg ttg gtg tgt       318
Arg Thr Val Phe Phe Trp Ala Pro Ile Met Lys Trp Gly Leu Val Cys
-25                 -20                 -15                 -10
gct gga ttg gct gat atg gcc aga cct gca gaa aaa ctt agc aca gct       366
```

74

```
Ala Gly Leu Ala Asp Met Ala Arg Pro Ala Glu Lys Leu Ser Thr Ala
              -5          1                    5
caa tct gct gtt ttg atg gct aca ggg ttt att tgg tca aga tac tca          414
Gln Ser Ala Val Leu Met Ala Thr Gly Phe Ile Trp Ser Arg Tyr Ser.
         10                      15                   20
ctt gta att att ccg aaa aat tgg agt ctg ttt gct gtt aat ttc ttt          462
Leu Val Ile Ile Pro Lys Asn Trp Ser Leu Phe Ala Val Asn Phe Phe
    25                       30                   35
gtg ggg gca gca gga gcc tct cag ctt ttt cgt att tgg aga tat aac          510
Val Gly Ala Ala Gly Ala Ser Gln Leu Phe Arg Ile Trp Arg Tyr Asn
40                   45                   50                   55
caa gaa cta aaa gct aaa gca cac aaa taaaagagtt cctgatcacc               557
Gln Glu Leu Lys Ala Lys Ala His Lys
                  60
tgaacaatct agatgtggac aaaaccattg ggacctagtt tattatttgg ttattgataa       617
agcaaagcta actgtgtgtt tagaaggcac tgtaactggt agctagttct tgattcaata       677
gaaaaatgca gcaaactttt aataacagtc tctctacatg acttaaggaa cttatctatg       737
gatattagta acatttttct accatttgtc cgtaataaaa catacttgct cgtaaaaaaa       797
aaaaaaa                                                                   804
```

<210> 63
<211> 792
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 194..253
<223> Von Heijne matrix
      score 12.3999996185303
      seq ALLLGALLGTAWA/RR
<220>
<221> polyA_signal
<222> 768..773
<220>
<221> polyA_site
<222> 780..792
<220>
<221> misc_feature
<222> 154..428
<223> homology
      id :R22491
      est
<220>
<221> misc_feature
<222> 104..160
<223> homology
      id :R22491
      est
<220>
<221> misc_feature
<222> 47..218
<223> homology
      id :AA136163
      est
<220>
<221> misc_feature
<222> 265..403
<223> homology

75

```
        id :AA136163
        est
<220>
<221> misc_feature
<222> 3..40
<223> homology
        id :AA136163
        est
<220>
<221> misc_feature
<222> 123..265
<223> homology
        id :N57089
        est
<220>
<221> misc_feature
<222> 47..127
<223> homology
        id :N57089
        est
<220>
<221> misc_feature
<222> 282..323
<223> homology
        id :N57089
        est
<220>
<221> misc_feature
<222> 128..403
<223> homology
        id :AA314970
        est
<220>
<221> misc_feature
<222> 138..403
<223> homology
        id :AA314807
        est
<220>
<221> misc_feature
<222> 164..403
<223> homology
        id :AA271811
        est
<220>
<221> misc_feature
<222> 163..385
<223> homology
        id :AA103053
        est
<220>
<221> misc_feature
<222> 154..403
<223> homology
        id :AA042016
        est
<220>
<221> misc_feature
```

```
<222> 2..250
<223> homology
      id :AA315322
      est
<220>
<221> misc_feature
<222> 154..403
<223> homology
      id :AA470189
      est
<220>
<221> misc_feature
<222> 217..403
<223> homology
      id :AA462839
      est
<220>
<221> misc_feature
<222> 154..403
<223> homology
      id :AA120322
      est
<220>
<221> misc_feature
<222> 163..403
<223> homology
      id :W71694
      est
<220>
<221> misc_feature
<222> 164..385
<223> homology
      id :AA250603
      est
<220>
<221> misc_feature
<222> 266..403
<223> homology
      id :AA036242
      est
<400> 63
aaggcggtcg ccgggacacc ccgtgtgtgg caggcggcga asgctctgga gaatcccgga        60
cagccctgct ccctgcagcc aggtgtagtt tcgggagcca ctggggccaa agtgagagtc       120
cagcggtctt ccagcgcttg ggccacggcg gcggccctgg gagcagaggt ggagcgaccc       180
cattacgcta aag atg aaa ggc tgg ggt tgg ctg gcc ctg ctt ctg ggg        229
            Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly
               -20               -15              -10
gcc ctg ctg gga acc gcc tgg gct cgg agg agc cgg gat ctc cac tgt       277
Ala Leu Leu Gly Thr Ala Trp Ala Arg Arg Ser Arg Asp Leu His Cys
         -5                    1             5
gga gca tgc agg gct ctg gtg gat gaa cta gaa tgg gaa att gcc cag       325
Gly Ala Cys Arg Ala Leu Val Asp Glu Leu Glu Trp Glu Ile Ala Gln
   10                   15                  20
gtg gac ccc aag aag acc att cag atg gga tcc ttc cgg atc aat cca       373
Val Asp Pro Lys Lys Thr Ile Gln Met Gly Ser Phe Arg Ile Asn Pro
25                   30                  35                  40
gat ggc agc cag tca gtg gtg gag gta act gtt act gkt tcc ccc aaa       421
Asp Gly Ser Gln Ser Val Val Glu Val Thr Val Thr Xaa Ser Pro Lys
```

77

```
                    45                      50                      55
aca aaa gta gct cac tct ggc ttt tgg atg aaa att cga ctg ctt aaa      469
Thr Lys Val Ala His Ser Gly Phe Trp Met Lys Ile Arg Leu Leu Lys
            60                      65                      70
aaa gga cct tgg tct taatagaaaa tgaagraaaa cagactcaga aaaaaagatt      524
Lys Gly Pro Trp Ser
            75
tbggctctgt ctcawtttgg aagaaggctg gcaggcttat tccccaatgc aactttgctt      584
cctggctgca aaccyttaat acytttgttt ctgctgtaga aatttgttag ccaaaacawg      644
ggagtcctga twcagcaacc ccttcttcca caatccacca tgactggttt ttaatgtamc      704
acttggggta tacatgcaaa accatccgtt cmaaaatctg aatycggagc ttaaaaattt      764
aaaaatgaaa aacchaaaaa aaaaaaaa                                         792
```

```
<210> 64
<211> 832
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 148..207
<223> Von Heijne matrix
      score 12.3999996185303
      seq ALLLGALLGTAWA/RR
<220>
<221> polyA_signal
<222> 789..794
<220>
<221> polyA_site
<222> 820..832
<220>
<221> misc_feature
<222> 258..553
<223> homology
      id :AA435303
      est
<220>
<221> misc_feature
<222> 117..219
<223> homology
      id :AA435303
      est
<220>
<221> misc_feature
<222> 552..645
<223> homology
      id :AA435303
      est
<220>
<221> misc_feature
<222> 217..258
<223> homology
      id :AA435303
      est
<220>
<221> misc_feature
<222> 258..553
<223> homology
      id :AA314807
```

```
        est
<220>
<221> misc_feature
<222> 92..258
<223> homology
        id :AA314807
        est
<220>
<221> misc_feature
<222> 258..554
<223> homology
        id :AA314970
        est
<220>
<221> misc_feature
<222> 82..258
<223> homology
        id :AA314970
        est
<220>
<221> misc_feature
<222> 258..553
<223> homology
        id :AA547310
        est
<220>
<221> misc_feature
<222> 119..258
<223> homology
        id :AA547310
        est
<220>
<221> misc_feature
<222> 359..553
<223> homology
        id :AA565602
        est
<220>
<221> misc_feature
<222> 552..683
<223> homology
        id :AA565602
        est
<220>
<221> misc_feature
<222> 684..751
<223> homology
        id :AA565602
        est
<220>
<221> misc_feature
<222> 742..783
<223> homology
        id :AA565602
        est
<220>
<221> misc_feature
<222> 364..553
```

<223> homology
      id :AA136094
      est   .
<220>
<221> misc_feature
<222> 552..683
<223> homology
      id :AA136094
      est
<220>
<221> misc_feature
<222> 684..751
<223> homology
      id :AA136094
      est
<220>
<221> misc_feature
<222> 258..461
<223> homology
      id :AA136163
      est
<220>
<221> misc_feature
<222> 2..172
<223> homology
      id :AA136163
      est
<220>
<221> misc_feature
<222> 216..258
<223> homology
      id :AA136163
      est
<220>
<221> misc_feature
<222> 743
<223> n=a, g, c or t
<400> 64
aggagaatcc cggacagccc tgctccctgc agccaggtgt agtttcggga gccactgggg   60
ccaaagtgag agtccagcgg tcttccagcg cttgggccac ggcggcggcc ctgggagcag   120
aggtggagcg accccattac gctaaag atg aaa ggc tgg ggt tgg ctg gcc ctg   174
                               Met Lys Gly Trp Gly Trp Leu Ala Leu
                               -20   -15
ctt ctg ggg gcc ctg ctg gga acc gcc tgg gct cgg agg agc cag gat   222
Leu Leu Gly Ala Leu Leu Gly Thr Ala Trp Ala Arg Arg Ser Gln Asp
    -10   -5   1   5
ctc cac tgt gga gca tgc agg gct ctg gtg gat gaa act aga atg gga   270
Leu His Cys Gly Ala Cys Arg Ala Leu Val Asp Glu Thr Arg Met Gly
                10   15   20
aat tgc cca ggt gga ccc caa gaa gac cat tca gat ggg atc ttt ccg   318
Asn Cys Pro Gly Gly Pro Gln Glu Asp His Ser Asp Gly Ile Phe Pro
            25   30   35
gat caa tcc aga tgg cag cca gtc agt ggt gga ggt gcc tta tgc ccg   366
Asp Gln Ser Arg Trp Gln Pro Val Ser Gly Gly Gly Ala Leu Cys Pro
            40   45   50
ctc aga ggc cca cct cac aga gct gct gga gga gat atg tgaccggatg   415
Leu Arg Gly Pro Pro His Arg Ala Ala Gly Gly Asp Met
    55   60   65
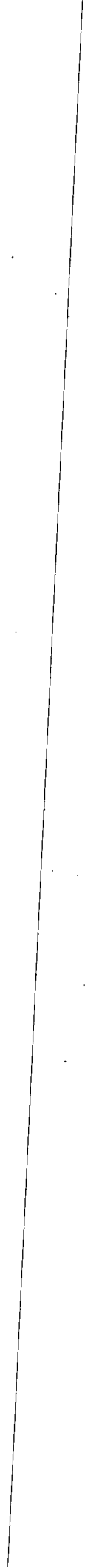
```
aaggagtatg gggaacagat tgatccttcc acccatcgca agaactacgt acgtgtagtg      475
ggccggaatg gagaatccag tgaactggac ctacaaggca tccgaatcga ctcagatatt      535
agcggcaccc tcaagbtttg cgtgtgggaa cattgtggag gaatacgagg atgaactcat      595
tgaattcttt tcccgagagg ctgacaatgt taaagacaaa ctttgcagta agcgaacaga      655
tctttgtgac catgccctgc acatatcggc atgatgagct atgaaccact ggagcagccc      715
acactggctt gatggatcac ccccaggnaa gggaaaatgg tggcaatgcc ttttatatat      775
tatgttttac tgaaattaac tgaaaaatat gaaaccaaaa gtscaaaaaa aaaaaaa        832
```

```
<210> 65
<211> 721
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 156..230
<223> Von Heijne matrix
      score 5
      seq MFAASLLAMCAGA/EV
<220>
<221> polyA_signal
<222> 706..711
<220>
<221> polyA_site
<222> 709..721
<220>
<221> misc_feature
<222> 351..688
<223> homology
       id :H98648
      est
<220>
<221> misc_feature
<222> 289..353
<223> homology
       id :H98648
      est
<220>
<221> misc_feature
<222> 274..641
<223> homology
       id :AA181022
      est
<220>
<221> misc_feature
<222> 255..286
<223> homology
       id :AA181022
      est
<220>
<221> misc_feature
<222> 242..641
<223> homology
       id :AA143192
      est
<220>
<221> misc_feature
<222> 261..646
<223> homology
```

81 210

```
        id :AA594850
        est
<220>
<221> misc_feature
<222> 165..474
<223> homology
        id :AA563681
       .est
<220>
<221> misc_feature
<222> 1..74
<223> homology
        id :AA563681
        est
<220>
<221> misc_feature
<222> 261..643
<223> homology
        id :AA287457
        est
<220>
<221> misc_feature
<222> 352..646
<223> homology
        id :N22567
        est
<220>
<221> misc_feature
<222> 299..354
<223> homology
        id :N22567
        est
<220>
<221> misc_feature
<222> 265..303
<223> homology
        id :N22567
        est
<220>
<221> misc_feature
<222> 30..165
<223> homology
        id :AA186657
        est
<220>
<221> misc_feature
<222> 270..349
<223> homology
        id :AA186657
        est
<220>
<221> misc_feature
<222> 213..261
<223> homology
        id :AA186657
        est
<220>
<221> misc_feature
```

```
<222> 165..214
<223> homology
      id :AA186657
      est
<220>
<221> misc_feature
<222> 346..387
<223> homology
      id :AA186657
      est
<220>
<221> misc_feature
<222> 52..400
<223> homology
      id :HSC1ED081
      est
<220>
<221> misc_feature
<222> 398..436
<223> homology
      id :HSC1ED081
      est
<220>
<221> misc_feature
<222> 171..316
<223> homology
      id :AA143136
      est
<400> 65
attttgggtc cggcctgctc gcmgtccgct ccgtccgccc ttagacctgt tgcccagcat      60
ccctgcagtt cgcggwacag tctctattag agcgcgtgta tagaggcaga kaggagtgaa     120
gtccacagtt cctctcctcc tagagcctgc cgacc atg ccc gcg ggc gtg ccc        173
                                        Met Pro Ala Gly Val Pro
                                        -25                -20
atg tcc acc tac ctg aaa atg ttc gca gcc agt ctc ctg gcc atg tgc      221
Met Ser Thr Tyr Leu Lys Met Phe Ala Ala Ser Leu Leu Ala Met Cys
             -15              -10                -5
gca ggg gca gaa gtg gtg cac agg tac tac cga ccg gac ctg aca ata      269
Ala Gly Ala Glu Val Val His Arg Tyr Tyr Arg Pro Asp Leu Thr Ile
             1               5                  10
cct gaa att cca cca aag cgt gga gaa ctc aaa acg gag ctt ttg gga      317
Pro Glu Ile Pro Pro Lys Arg Gly Glu Leu Lys Thr Glu Leu Leu Gly
     15              20              25
ctg aaa gaa aga aaa cac aaa cct caa gtt tct caa cag gag gaa ctt      365
Leu Lys Glu Arg Lys His Lys Pro Gln Val Ser Gln Gln Glu Glu Leu
30                  35              40                  45
aaa taactatgcc aagaattctg tgaataatat aagtcttaaa tatgtatttc           418
Lys
ttaatttatt gcatcaaact acttgtcctt aagcacttag tctaatgcta actgcaagag     478
gaggtgctca gtggatgttt agccgatacg ttgaaattta attacggttt gattgatatt     538
tcttgaaaac tgccaaagca catatcatca aaccatttca tgaatatggt ttggaagatg     598
tttagtcttg aatataacgc gaaatagaat atttgtaagt ctactatatg ggttgtcttt     658
atttcatata aattaagaaa ttatttaaaa actatgaact aggtttcatt aaaaaaaaaa     718
gaa                                                                   721

<210> 66
<211> 531
<212> DNA
```

83 212

```
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 272..397
<223> Von Heijne matrix
      score 4.59999990463257
      seq RIPSLPGSPVCWA/WP
<220>
<221> polyA_signal
<222> 503..508
<220>
<221> polyA_site
<222> 518..531
<220>
<221> misc_feature
<222> 235..517
<223> homology
      id :AA524403
      est
<220>
<221> misc_feature
<222> 52..208
<223> homology
      id :AA524403
      est
<220>
<221> misc_feature
<222> 259..517
<223> homology
      id :N93600
      est
<220>
<221> misc_feature
<222> 85..207
<223> homology
      id :N93600
      est
<220>
<221> misc_feature
<222> 353..517
<223> homology
      id :AA594610
      est
<220>
<221> misc_feature
<222> 258..363
<223> homology
      id :AA594610
      est
<220>
<221> misc_feature
<222> 105..207
<223> homology
      id :AA594610
      est
<220>
<221> misc_feature
<222> 202..517
```

213

```
<223> homology
      id :AA074748
      est
<220>
<221> misc_feature
<222> 116..153
<223> homology
      id :AA074748
      est
<220>
<221> misc_feature
<222> 167..202
<223> homology
      id :AA074748
      est
<220>
<221> misc_feature
<222> 258..517
<223> homology
      id :N93603
      est
<220>
<221> misc_feature
<222> 208..251
<223> homology
      id :N93603
      est
<220>
<221> misc_feature
<222> 163..202
<223> homology
      id :N93603
      est
<220>
<221> misc_feature
<222> 90..125
<223> homology
      id :N93603
      est
<220>
<221> misc_feature
<222> 125..363
<223> homology
      id :HSPD04938
      est
<220>
<221> misc_feature
<222> 353..517
<223> homology
      id :HSPD04938
      est
<220>
<221> misc_feature
<222> 28..227
<223> homology
      id :AA074804
      est
<220>
```

```
<221> misc_feature
<222> 265..310
<223> homology
      id :AA074804
      est
<220>
<221> misc_feature
<222> 227..263
<223> homology
      id :AA074804
      est
<220>
<221> misc_feature
<222> 352..385
<223> homology
      id :AA074804
      est
<400> 66
aaaaggaaag aggtysggag cgctcgcgag atctcggacc acccaacctg aaaggtgctt      60
aggaagttga aaggcccaga ggaggcctcc gggcaaatgg ccggagctgg accgaccatg     120
ctgctacgag aagagaatgg ctgttgcagt cggcgtcaga gcagctccag tgccggggat     180
tcggacggag agcgcgagga ctcggcggct gagcgcgccc gacagcagct agaggcgctg     240
ctcaacaaga ctatgcgcat cgcatgaca g atg gac gga cac tgg tcg gct         292
                                   Met Asp Gly His Trp Ser Ala
                                                         -40
gct ttc tct gca ctg acc gtg act gca atg tca tcc tgg gct cgg cgc       340
Ala Phe Ser Ala Leu Thr Val Thr Ala Met Ser Ser Trp Ala Arg Arg
-35                     -30                 -25                 -20
agg agt tcc tca agc cgt cgg att cct tct ctg ccg ggg agc ccc gtg       388
Arg Ser Ser Ser Ser Arg Arg Ile Pro Ser Leu Pro Gly Ser Pro Val
                -15                 -10                 -5
tgc tgg gcc tgg cca tgg tac ccg gac acc aca tcg ttt cca ttg agg       436
Cys Trp Ala Trp Pro Trp Tyr Pro Asp Thr Thr Ser Phe Pro Leu Arg
            1               5                   10
tgc aga ggg aga gtc tgaccgggcc tccgtatctc tgaccacgat ggcgcttacc      491
Cys Arg Gly Arg Val
    15
tttcagactt cattaaactt atgaccaaaa aaaaaaaaa                           531

<210> 67
<211> 783
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 381..629
<223> Von Heijne matrix
      score 8.60000038146973
      seq LELLTSCSPPASA/SQ
<220>
<221> polyA_signal
<222> 736..741
<220>
<221> polyA_site
<222> 770..783
<220>
<221> misc_feature
<222> 207..263
```

<223> homology
        id :AA357230
        est
<400> 67
agggacttcc ggcctcgctg gcgtggacgt ttgtggtggg gcgtgttggt ccgcgctctc   60
agaactgtgc tgggaaggat ggtagggcga ctggggctca cctccgcacc gttgtaggac   120
ccggggtagg gtttttgagcc cgtgggagct gccccacgcg gcctcgtcct gccaacggtc   180
ggatggcgga gacgaaggac gcagcgcaga tgttggtgac cttcaaggat gtggctgtga   240
cctttacccg ggaggagtgg agacagctgg acctggccca gaggaccctg taccgagagg   300
tgatcgggtt cccaaaccag agttggtcca cctgctagag catgggcagg agctgtggat   360
agtgaagaga ggcctctcac atg cta cct gtg cag agt ttc act ctt gtt gcc   413
                        Met Leu Pro Val Gln Ser Phe Thr Leu Val Ala
                                         -80   -75
cag gct gga gtg cag tgg cgc cat ctc agc tca ctg caa ctt ctg cct   461
Gln Ala Gly Val Gln Trp Arg His Leu Ser Ser Leu Gln Leu Leu Pro
        -70   -65   -60
ccc gag ttc aag gga ttc tcc tgc ctc agc ctc ccg agt agc tgg gat   509
Pro Glu Phe Lys Gly Phe Ser Cys Leu Ser Leu Pro Ser Ser Trp Asp
        -55   -50   -45
tac agg cgc cca cca cca tgc ccg gct ggt ttt ttt gta ttt tta gta   557
Tyr Arg Arg Pro Pro Pro Cys Pro Ala Gly Phe Phe Val Phe Leu Val
-40   -35   -30   -25
gag acg ggg ctt cac cat gtt ggc cag gct ggt ctt gaa ctc ttg acc   605
Glu Thr Gly Leu His His Val Gly Gln Ala Gly Leu Glu Leu Leu Thr
                -20   -15   -10
tca tgt agt cca ccc gcc tct gcc tcc caa agt gct gcg att aca ggc   653
Ser Cys Ser Pro Pro Ala Ser Ala Ser Gln Ser Ala Ala Ile Thr Gly
                -5   1   5
gtg agc cac gtg ccc ggc aaa aaa aaa ctg ctt aag gtt gaa aag aaa   701
Val Ser His Val Pro Gly Lys Lys Lys Leu Leu Lys Val Glu Lys Lys
        10   15   20
aat tta aga aaw ttg ctg acg gra ata aaa acy taataaaact accacccgaa   754
Asn Leu Arg Xaa Leu Leu Thr Xaa Ile Lys Thr
25   30   35
ggaatgaaaa aaccaaaaaa aaaaaaaaa   783

<210> 68
<211> 996
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 140..205
<223> Von Heijne matrix
        score 5.90000009536743
        seq IILGCLALFLLLQ/RK
<220>
<221> polyA_signal
<222> 965..970
<220>
<221> polyA_site
<222> 984..996
<220>
<221> misc_feature
<222> 676..959
<223> homology
        id :AA399103
        est

<220>
<221> misc_feature
<222> 609..679
<223> homology
        id :AA399103
        est
<220>
<221> misc_feature
<222> 225..433
<223> homology
        id :AA398040
        est
<220>
<221> misc_feature
<222> 433..563
<223> homology
        id :AA398040
        est
<400> 68

```
aacagttacg aaggagagct gcaaaagttg cagcagaaag gttgggagtc ccgacaggtt      60
ccgtagccca cagaaaagaa gcaagggacg gcaggactgt ttcacacttt tctgcttctg     120
gaaggtgctg acaaaaac atg gaa cta att tcc cca aca gtg att ata atc       172
                    Met Glu Leu Ile Ser Pro Thr Val Ile Ile Ile
                                  -20                 -15
ctg ggt tgc ctt gct ctg ttc tta ctc ctt cag cgg aag aat ttg cgc       220
Leu Gly Cys Leu Ala Leu Phe Leu Leu Leu Gln Arg Lys Asn Leu Arg
     -10                  -5                      1                5
aga ccc ccg tgc atc aag ggc tgg att cct tgg att gga gtt gga ttt       268
Arg Pro Pro Cys Ile Lys Gly Trp Ile Pro Trp Ile Gly Val Gly Phe
              10                  15                  20
gak ttt ggg aaa gcc cct cta gaa ttt ata gag aaa gca aga atc aag       316
Xaa Phe Gly Lys Ala Pro Leu Glu Phe Ile Glu Lys Ala Arg Ile Lys
              25                  30                  35
gta tgt ggt cgt ggc ava cgg ggt ctc cag agg aga caa tgc ttt ctt       364
Val Cys Gly Arg Gly Xaa Arg Gly Leu Gln Arg Arg Gln Cys Phe Leu
              40                  45                  50
ttt taaactttct ttcattgact cttaagtgca gggctagaac acggggaaca           417
Phe
tacctgcttg cctcaaacta aaggatctag tcmtytctga aktcctctac tsacrrttra     477
caacaatatc ctgtgcaaaa ttttgcgaaa gaaatgaaat acaattgcmg cgtgcatcga     537
catttttgga agtagagatt aacyyttcgt atttttactt cmtcgaagtt aagttccaaa     597
tgtgtatgtg ttaagtaaat gttttcagta aytgggaaag ataaagtgta atccaattta     657
agtttgtgaa aatgagtaat tccgtatcca aaytggagtt aacaccaaag tattgtacaa     717
attgcttgca cagttggtcc gtacacaata gacaggctyt gtattttag ctgacgttgt      777
tatttgatga tgatgtactc cattttcamt acggcccgaa gagamtagta atcctccttg     837
tagtagatgt ttttgtcttg aaagtatctt ttaaatgtyt gagcacttta aggaacagac     897
ccttattaat gtyttttaag ttttattcaa tttccagtca caaatatttt atggtatttg     957
attgtytaat aaatttgtat gatattaaaa aaaaaaaaa                            996
```

<210> 69
<211> 657
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 183..338
<223> Von Heijne matrix
        score 3.79999995231628

```
          seq VMLETCGLLVSLG/QS
<220>
<221> polyA_signal
<222> 620..625
<220>
<221> polyA_site
<222> 644..657
<220>
<221> misc_feature
<222> 207..263
<223> homology
      id :AA357230
      est
<400> 69
agggacttcc ggcctcgctg gcgtggacgt ttgtggtggg gcgtgttggt ccgcgctctc    60
agaactgtgc tgggaaggat ggtagggcga ctggggctca cctccgcacc gttgtaggac   120
ccggggtagg gttttgagcc cgtgggagct gccccacgcg gcctcgtcct gccaacggtc   180
gg atg gcg gag acg aag gac gca gcg cag atg ttg gtg acc ttc aag       227
   Met Ala Glu Thr Lys Asp Ala Ala Gln Met Leu Val Thr Phe Lys
           -50                 -45                 -40
gat gtg gct gtg acc ttt acc cgg gag gag tgg aga cag ctg gac ctg      275
Asp Val Ala Val Thr Phe Thr Arg Glu Glu Trp Arg Gln Leu Asp Leu
        -35                 -30                 -25
gcc cag agg acc ctg tac cga gag gtg atg ctg gag acc tgt ggg ctt      323
Ala Gln Arg Thr Leu Tyr Arg Glu Val Met Leu Glu Thr Cys Gly Leu
        -20                 -15                 -10
ctg gtt tca cta ggg caa agc att tgg ctg cat ata aca gaa aac cag      371
Leu Val Ser Leu Gly Gln Ser Ile Trp Leu His Ile Thr Glu Asn Gln
-5                  1                   5                   10
atc aaa ctg gct tca cct gga agg aaa ttc act aac tcg cct gat gag      419
Ile Lys Leu Ala Ser Pro Gly Arg Lys Phe Thr Asn Ser Pro Asp Glu
           15                  20                  25
aag cct gag gtg tgg ttg gct cca ggc ctg ttc ggt gcc gca gcc cag      467
Lys Pro Glu Val Trp Leu Ala Pro Gly Leu Phe Gly Ala Ala Ala Gln
           30                  35                  40
tgacgccatc aaggatgtct tggttctctg ttccttcttc ttggttcagg cttctggatt   527
gtcctcaggc tggctcctca tagggatgct gggtgctgca gccttgactg gggcagcagg   587
cccccatggt tcaatccatc ctcccacctt ggaataaatg ctttcttttc acaatgagaa   647
aaaaaaaaaa                                                           657

<210> 70
<211> 416
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 140..205
<223> Von Heijne matrix
      score 5.90000009536743
      seq IILGCLALFLLLQ/RK
<220>
<221> polyA_signal
<222> 383..388
<220>
<221> polyA_site
<222> 405..416
<220>
<221> misc_feature
```

89

<222> 225..316
<223> homology
        id :AA398040
        est
<400> 70
aacagttacg aaggagagct gcaaaagttg cagcagaaag gttgggagtc ccgacaggtt   60
ccgtagccca cagaaaagaa gcaagggacg gcaggactgt ttcacacttt tctgcttctg   120
gaaggtgctg acaaaaac atg gaa cta att tcc cca aca gtg att ata atc   172
                    Met Glu Leu Ile Ser Pro Thr Val Ile Ile Ile
                          -20   -15
ctg ggt tgc ctt gct ctg ttc tta ctc ctt cag cgg aag aat ttg cgc   220
Leu Gly Cys Leu Ala Leu Phe Leu Leu Leu Gln Arg Lys Asn Leu Arg
      -10   -5   1   5
aga ccc ccg tgc atc aag ggc tgg att cct tgg att gga gtt gga ttt   268
Arg Pro Pro Cys Ile Lys Gly Trp Ile Pro Trp Ile Gly Val Gly Phe
                  10   15   20
gag ttt ggg aaa gcc cct cta gaa ttt ata gag aaa gca aga atc aag   316
Glu Phe Gly Lys Ala Pro Leu Glu Phe Ile Glu Lys Ala Arg Ile Lys
          25   30   35
tat gga cca ata ttt aca gtc ttt gct atg gga aac cga atg acc ttt   364
Tyr Gly Pro Ile Phe Thr Val Phe Ala Met Gly Asn Arg Met Thr Phe
          40   45   50
gtt act gaa gaa gga agg aat taatgtgttt ctaaatcca aaaaaaaaaa a   416
Val Thr Glu Glu Gly Arg Asn
      55   60

<210> 71
<211> 543
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 129..176
<223> Von Heijne matrix
        score 4.80000019073486
        seq SLFIYIFLTCSNT/SP
<220>
<221> polyA_signal
<222> 513..518
<220>
<221> polyA_site
<222> 530..543
<220>
<221> misc_feature
<222> 264..500
<223> homology
        id :AA534039
        est
<220>
<221> misc_feature
<222> 205..315
<223> homology
        id :T82645
        est
<220>
<221> misc_feature
<222> 295..382
<223> homology

90

```
        id :T82645
        est
<220>
<221> misc_feature
<222> 375..405
<223> homology
        id :T82645
        est
<220>
<221> misc_feature
<222> 50
<223> n=a, g, c or t
<400> 71
actgtcccat tcctccccct acaacacaca cacctttcag gcagggasgn gatgagcttc       60
cagccccaag agtggaggct gccacatcct aacatasgta tctattgaaa aggaagcagt      120
gtgtatct atg att ata tct ctg ttc atc tat ata ttt ttg aca tgt agc       170
         Met Ile Ile Ser Leu Phe Ile Tyr Ile Phe Leu Thr Cys Ser
              -15                 -10                 -5
aac acc tct cca tct tat caa gga act caa ctc ggt ctg ggt ctc ccc       218
Asn Thr Ser Pro Ser Tyr Gln Gly Thr Gln Leu Gly Leu Gly Leu Pro
         1     .           5                   10
agt gcc cag tgg tgg cct ttg aca ggt agg agg atg cag tgc tgc agg       266
Ser Ala Gln Trp Trp Pro Leu Thr Gly Arg Arg Met Gln Cys Cys Arg
15                  20                  25                  30
cta ttt tgt ttt ttg tta caa aac tgt ctt ttc cct ttt ccc ctc cac       314
Leu Phe Cys Phe Leu Leu Gln Asn Cys Leu Phe Pro Phe Pro Leu His
              35                  40                  45
ctg att cag cat gat ccc tgt gag ctg gtt ctc aca atc tcc tgg gac       362
Leu Ile Gln His Asp Pro Cys Glu Leu Val Leu Thr Ile Ser Trp Asp
              50                  55                  60
tgg gct gag gca ggg gct tcg ctc tat tct ccc taaccatact gtcttccttt      415
Trp Ala Glu Ala Gly Ala Ser Leu Tyr Ser Pro
              65                  70
cccccttgcc acttagcagt tatcccccca gctatgcctt ctccctccct cccttgccct      475
ggcatatatt gtgccttatt tatgctgcaa atataacatt aaactatcaa gtgaaaaaaa      535
aaaaaaaa                                                                543

<210> 72
<211> 605
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 285..341
<223> Von Heijne matrix
        score 5.59999990463257
        seq PTLCVSSSPALWA/AS
<220>
<221> polyA_signal
<222> 575..580
<220>
<221> polyA_site
<222> 592..605
<220>
<221> misc_feature
<222> 53..296
<223> homology
        id :W07033
```

```
        est
<220>
<221> misc_feature
<222> 348..432
<223> homology
        id :W07033
        est
<220>
<221> misc_feature
<222> 435..497
<223> homology
        id :W07033
        est
<220>
<221> misc_feature
<222> 293..337
<223> homology
        id :W07033
        est
<220>
<221> misc_feature
<222> 521..560
<223> homology
        id :W07033
        est
<220>
<221> misc_feature
<222> 489..520
<223> homology
        id :W07033
        est
<220>
<221> misc_feature
<222> 15..337
<223> homology
        id :AA151004
        est
<220>
<221> misc_feature
<222> 348..412
<223> homology
        id :AA151004
        est
<220>
<221> misc_feature
<222> 434..485
<223> homology
        id :AA151004
        est
<220>
<221> misc_feature
<222> 83..324
<223> homology
        id :AA476506
        est
<220>
<221> misc_feature
<222> 347..560
```
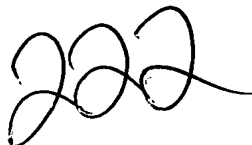
```
<223> homology
      id :AA476506
      est
<220>
<221> misc_feature
<222> 16..347
<223> homology
      id :W56567
      est
<220>
<221> misc_feature
<222> 350..405
<223> homology
      id :W56567
      est
<220>
<221> misc_feature
<222> 433..470
<223> homology
      id :W56567
      est
<220>
<221> misc_feature
<222> 15..296
<223> homology
      id :AA147584
      est
<220>
<221> misc_feature
<222> 348..421
<223> homology
      id :AA147584
      est
<220>
<221> misc_feature
<222> 293..337
<223> homology
      id :AA147584
      est
<220>
<221> misc_feature
<222> 419..453
<223> homology
      id :AA147584
      est
<220>
<221> misc_feature
<222> 2..338
<223> homology
      id :AA281959
      est
<220>
<221> misc_feature
<222> 350..432
<223> homology
      id :AA281959
      est
<400> 72
```

```
aacgcctwta agacagcgga actaagaaaa gaagaggcct gtggacagaa caatcatgtc      60
tgactccctg gtggtgtgcg aggtagaccc agagctaaca gaaaagctga kgaaattccg     120
cttccgaaaa gagacagaca atgcagccat cataatgaag gtggacaaag accggcagat     180
ggtggtgctg gaggaagaat ttcagaacat ttccccagag gagctcaaaa tggagttgcc     240
ggagagacag cccaggttcg tggtttacag ctacaagtac gtgc atg acg atg gcc     296
                                                    Met Thr Met Ala

gag tgt cct acc ctt tgt gtt tca tct tct cca gcc ctg tgg gct gca     344
Glu Cys Pro Thr Leu Cys Val Ser Ser Ser Pro Ala Leu Trp Ala Ala
-15                     -10                 -5                    1
agc gaa aca aca gat gat gta tgc agg gag taaaaacagg ctggtgcaga        394
Ser Glu Thr Thr Asp Asp Val Cys Arg Glu
            5                   10
cagcagagct cacaaaggtg ttcgaaatcc gcaccactga tgacctcact gaggcctggc     454
tccaagaaaa gttgtctttc tttcgttgat ctctgggctg gggactgaat tcctgatgtc     514
tgagtcctca aggtgactgg ggacttggaa cccctaggac ctgaacaacc aaggacttta     574
aataaatttt aaaatgcaaa aaaaaaaaa a                                     605
```

<210> 73
<211> 864
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 136..444
<223> Von Heijne matrix
      score 4.90000009536743
      seq VYAFLGLTAPSGS/KE
<220>
<221> polyA_signal
<222> 835..840
<220>
<221> polyA_site
<222> 851..864
<220>
<221> misc_feature
<222> 222..456
<223> homology
      id :AA136758
      est
<220>
<221> misc_feature
<222> 557..648
<223> homology
      id :AA136758
      est
<220>
<221> misc_feature
<222> 501..571
<223> homology
      id :AA136758
      est
<220>
<221> misc_feature
<222> 130..456
<223> homology
      id :AA393612
      est
<220>

94

```
<221> misc_feature
<222> 88..130
<223> homology
        id :AA393612
        est
<220>
<221> misc_feature
<222> 501..538
<223> homology
        id :AA393612
        est
<220>
<221> misc_feature
<222> 130..458
<223> homology
        id :R59039
        est
<220>
<221> misc_feature
<222> 71..130
<223> homology
        id :R59039
        est
<220>
<221> misc_feature
<222> 557..716
<223> homology
        id :W48624
        est
<220>
<221> misc_feature
<222> 365..456
<223> homology
        id :W48624
        est
<220>
<221> misc_feature
<222> 501..571
<223> homology
        id :W48624
        est
<220>
<221> misc_feature
<222> 716..751
<223> homology
        id :W48624
        est
<220>
<221> misc_feature
<222> 222..458
<223> homology
        id :AA136810
        est
<220>
<221> misc_feature
<222> 501..581
<223> homology
        id :AA136810
```

```
            est
<220>
<221> misc_feature
<222> 587..668
<223> homology
        id :AA136810
        est
<220>
<221> misc_feature
<222> 130..419
<223> homology
        id :T35647
        est
<220>
<221> misc_feature
<222> 59..130
<223> homology
        id :T35647
        est
<220>
<221> misc_feature
<222> 557..852
<223> homology
        id :HUM093F06A
        est
<220>
<221> misc_feature
<222> 501..571
<223> homology
        id :HUM093F06A
        est
<220>
<221> misc_feature
<222> 130..384
<223> homology
        id :T35666
        est
<400> 73
aaagttctcc ttccaccttc ccccaccctt ctctgccaac cgctgtttca gcccctagct        60
ggattccagc cattgctgca gctgctccac agcccttttc aggacccaaa caaccgcagc       120
cgctgttccc caggr atg gtg atc cgt gta tat att gca tct tcc tct ggc       171
              Met Val Ile Arg Val Tyr Ile Ala Ser Ser Ser Gly
                        -100                    -95
tct aca gcg att aag aag aaa caa caa gat gtg ctt ggt ttc cta gaa       219
Ser Thr Ala Ile Lys Lys Lys Gln Gln Asp Val Leu Gly Phe Leu Glu
    -90               -85                   -80
gcc aac aaa ata gga ttt gaa gaa aaa gat att gca gcc aat gaa gag       267
Ala Asn Lys Ile Gly Phe Glu Glu Lys Asp Ile Ala Ala Asn Glu Glu
-75               -70                   -65                   -60
aat cgg aag tgg atg aga gaa aat gta cct gaa aat agt cga cca gcc       315
Asn Arg Lys Trp Met Arg Glu Asn Val Pro Glu Asn Ser Arg Pro Ala
              -55                   -50                   -45
aca ggt aac ccc ctg cca cct cag att ttc aat gaa agc cag tat cgc       363
Thr Gly Asn Pro Leu Pro Pro Gln Ile Phe Asn Glu Ser Gln Tyr Arg
              -40                   -35                   -30
ggg gac tat gat gcc ttc ttt gaa gcc aga gaa aat aat gca gtg tat       411
Gly Asp Tyr Asp Ala Phe Phe Glu Ala Arg Glu Asn Asn Ala Val Tyr
          -25                   -20                   -15
```
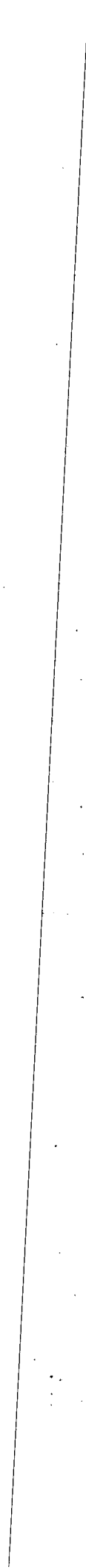
```
gcc ttc tta ggc ttg aca gcc cca tct ggt tca aag gaa gca gga agg          459
Ala Phe Leu Gly Leu Thr Ala Pro Ser Gly Ser Lys Glu Ala Gly Arg
    -10                 -5                  1                   5
tgc aag caa agc agc aag cca tgaaccttga gcactgtgct tttaagcatc            510
Cys Lys Gln Ser Ser Lys Pro
                10
ctgaaaaatg agtctccatt gctttttataa aatagcagaa ttagctttgc sttcaaaaga      570
aataggstta atgttgaaat aatagattag ttgggttttc acatgcaaac amtcaaaatg       630
aatacaaaat taaaatttga acattatggt gattatggtg aggagaatgg gatattaaca       690
taaaattata ttaataagta gatatygtag aaatagtgtt gttacctgcc aagccatcct       750
gtatacacca atgattttac aaagaaaaca cccttccctc cttytgccat tamtatggca       810
acctaagtgt atytgcagct ttacattaaa aaggagaaag agaaaaaaaa aaaa             864


<210> 74
<211> 1033
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 200..427
<223> Von Heijne matrix
      score 4.69999980926514
      seq LIVYLWVVSFIAS/SS
<220>
<221> polyA_signal
<222> 1001..1006
<220>
<221> polyA_site
<222> 1022..1033
<220>
<221> misc_feature
<222> 55..406
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 397..487
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 527..584
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 482..531
<223> homology
      id :AA056667
      est
<220>
<221> misc_feature
<222> 581..634
<223> homology
      id :AA056667
```

97

```
        est
<220>
<221> misc_feature
<222> 397..700
<223> homology
        id :AA044187
        est
<220>
<221> misc_feature
<222> 222..406
<223> homology
        id :AA044187
        est
<220>
<221> misc_feature
<222> 693..748
<223> homology
        id :AA044187
        est
<220>
<221> misc_feature
<222> 68..406
<223> homology
        id :AA131958
        est
<220>
<221> misc_feature
<222> 397..517
<223> homology
        id :AA131958
        est
<220>
<221> misc_feature
<222> 510..558
<223> homology
        id :AA131958
        est
<220>
<221> misc_feature
<222> 77..531
<223> homology
        id :W95957
        est
<220>
<221> misc_feature
<222> 527..558
<223> homology
        id :W95957
        est
<220>
<221> misc_feature
<222> 397..586
<223> homology
        id :AA041216
        est
<220>
<221> misc_feature
<222> 286..406
```
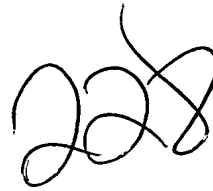
```
<223> homology
      id :AA041216
      est
<220>
<221> misc_feature
<222> 582..700
<223> homology
      id :AA041216
      est
<220>
<221> misc_feature
<222> 77..406
<223> homology
      id :W95790
      est
<220>
<221> misc_feature
<222> 397..539
<223> homology
      id :W95790
      est
<220>
<221> misc_feature
<222> 474..760
<223> homology
      id :AA461134
      est
<220>
<221> misc_feature
<222> 788..940
<223> homology
      id :AA461134
      est
<400> 74
aagacgaggt catgaatcat gtgacggtgg cttgaggagg aacctgtctt taaagctgtc      60
cctgaagtga cagcggagag aaccaggcag cccagaaacc ccaggcgtgg agattgatcc     120
tgcgagagaa gggggttcat catggcggat gacctaaagc gattcttgta taaaaagtta     180
ccaagtgttg aagggctcc atg cca ttg ttg tgt cag ata gag atg gag tac     232
                     Met Pro Leu Leu Cys Gln Ile Glu Met Glu Tyr
                         -75                 -70
ctg tta tta aag tgg caa atg aca atg ctc cag agc atg ctt tgc gac     280
Leu Leu Leu Lys Trp Gln Met Thr Met Leu Gln Ser Met Leu Cys Asp
-65                 -60                 -55                 -50
ctg gtt tct tat cca ctt ttg ccc ttg caa cag acc aag gaa gca aac     328
Leu Val Ser Tyr Pro Leu Leu Pro Leu Gln Gln Thr Lys Glu Ala Asn
                -45                 -40                 -35
ttg gac ttt cca aaa ata aaa gta tca tct gtt act ata aca cct acc     376
Leu Asp Phe Pro Lys Ile Lys Val Ser Ser Val Thr Ile Thr Pro Thr
            -30                 -25                 -20
agg tgg ttc aat tta atc gtt tac ctt tgg gtg gtg agt ttc ata gcc     424
Arg Trp Phe Asn Leu Ile Val Tyr Leu Trp Val Val Ser Phe Ile Ala
        -15                 -10                 -5
agc agc agt gcc aat aca gga cta att gtc agc cta gaa aag gaa ctt     472
Ser Ser Ser Ala Asn Thr Gly Leu Ile Val Ser Leu Glu Lys Glu Leu
  1                 5                   10                  15
gct cca ttg ttt gaa gaa ctg aga caa gtt gtg gaa gtt tct            514
Ala Pro Leu Phe Glu Glu Leu Arg Gln Val Val Glu Val Ser
            20                  25
```

99

```
taatctgaca gtggtttcag tgtgtacctt atcttcatta taacaacaca atatcaatcc    574
agcaatcttt agactacaat aatactttta tccatgtgct caagaaaggg cccctttttc    634
caacttatac taaagagcta gcatatagat gtaatttata gatagatcag ttgctatatt    694
ttctggtgta gggtctttct tatttagtga gatctaggga taccacagaa atggttcagt    754
ctatcaacag ctcccatgga gttagtctgg tcacagatat ggatgagaga ttytattcag    814
tggatcagaa tcaaactggt acattgatcc acttgagccg ttaagtgctg ccaattgtac    874
aatatgccca ggcttgcaga ataaagccaa ctttttattg tgaataataa taaggacata    934
tttttyttca gattatgttt tatttytttg cattgagtga ggaacataaa atggcttggt    994
aaaagtaata aaatcagtac aatcactaaa aaaaaaaaa                          1033
```

<210> 75
<211> 499
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 68..133
<223> Von Heijne matrix
        score 9.80000019073486
        seq LVVFCLALQLVPG/SP
<220>
<221> polyA_signal
<222> 472..477
<220>
<221> polyA_site
<222> 490..499
<400> 75

```
aaacagcagt gcctggtcaa acccagcaac ccttggccag aacttactca cccatcccac    60
tgacacc atg aag cct gtg ctg cct ctc cag ttc ctg gtg gtg ttc tgc    109
        Met Lys Pro Val Leu Pro Leu Gln Phe Leu Val Val Phe Cys
            -20                   -15                   -10
cta gca ctg cag ctg gtg cct ggg agt ccc aag cag cgt gtt ctg aag    157
Leu Ala Leu Gln Leu Val Pro Gly Ser Pro Lys Gln Arg Val Leu Lys
        -5                    1                   5
tat atc ttg gaa cct cca ccc tgc ata tca gca cct gaa aac tgt act    205
Tyr Ile Leu Glu Pro Pro Pro Cys Ile Ser Ala Pro Glu Asn Cys Thr
    10                  15                  20
cac ctg tgt aca atg cag gaa gat tgc gag aaa gga ttt cag tgc tgt    253
His Leu Cys Thr Met Gln Glu Asp Cys Glu Lys Gly Phe Gln Cys Cys
25                  30                  35                  40
tcc tcc ttc tgt ggg ata gtc tgt tca tca gaa aca ttt caa aag cgc    301
Ser Ser Phe Cys Gly Ile Val Cys Ser Ser Glu Thr Phe Gln Lys Arg
                45                  50                  55
aac aga atc aaa cac aag ggc tca gaa gtc atc atg cct gcc aac        346
Asn Arg Ile Lys His Lys Gly Ser Glu Val Ile Met Pro Ala Asn
            60                  65                  70
tgaggcatat ttcctagatc attttgcctc tacgatgttt tttcttggtc cacctttagg    406
aaggtattga gaagcaagaa actggaggcc caatatctaa cctgcaaatc gtttttgagt    466
ttggcaataa aggctaatct accaaaaaaa aaa                                499
```

<210> 76
<211> 978
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 274..399
<223> Von Heijne matrix

100

```
        score 5.19999980926514
        seq LLFDLVCHEFCQS/DD
<220>
<221> polyA_signal
<222> 943..948
<220>
<221> polyA_site
<222> 966..978
<220>
<221> misc_feature
<222> 335..518
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 225..274
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 812..861
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 186..224
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 708..748
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 276..314
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 146..176
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
<222> 879..909
<223> homology
        id :AA206225
        est
<220>
<221> misc_feature
```

```
<222> 182..518
<223> homology
      id :C15003
      est
<220>
<221> misc_feature
<222> 708..748
<223> homology
      id :C15003
      est
<220>
<221> misc_feature
<222> 182..517
<223> homology
      id :HUM407E11B
      est
<220>
<221> misc_feature
<222> 170..202
<223> homology
      id :AA544037
      est
<220>
<221> misc_feature
<222> 517..595
<223> homology
      id :HUM00TW170
      est
<220>
<221> misc_feature
<222> 596..665
<223> homology
      id :HUM00TW170
      est
<220>
<221> misc_feature
<222> 697..748
<223> homology
      id :HUM00TW170
      est
<220>
<221> misc_feature
<222> 805..861
<223> homology
      id :HUM00TW170
      est
<220>
<221> misc_feature
<222> 212..369
<223> homology
      id :HUM169E08B
      est
<220>
<221> misc_feature
<222> 406..493
<223> homology
      id :HUM169E08B
      est
```

<220>
<221> misc_feature
<222> 542..595
<223> homology
       id :HUM00TW112
       est
<220>
<221> misc_feature
<222> 697..748
<223> homology
       id :HUM00TW112
       est
<400> 76
accaggaaca tccagctatt tatgatagca tttgcttcat tatgtcaagt tcaacaaatg   60
ttgacttgct ggtgaaggtg ggggaggttg tggacaagct ctttgatttg gatgagaaac   120
taatgttaag aatgggtcag aaatggggct gctcagcctc tggaccaacc ccaggaagag   180
tctgaagagc agccagtgtt tcggcttgtg ccctgtatac ttgaagctgc caaacaagta   240
cgttctgaaa atccagaatg gcttgatgtt tac atg cac att tta caa ctg ctt   294
                                  Met His Ile Leu Gln Leu Leu
                                                              -40
act aca gtg gat gat gga att caa gca att gta cat tgt cct gac act   342
Thr Thr Val Asp Asp Gly Ile Gln Ala Ile Val His Cys Pro Asp Thr
-35   -30   -25   -20
gga aaa gac att tgg aat tta ctt ttt gac ctg gtc tgc cat gaa ttc   390
Gly Lys Asp Ile Trp Asn Leu Leu Phe Asp Leu Val Cys His Glu Phe
                    -15   -10   -5
tgc cag tct gat gat cca gcc atc att ctt caa gaa cag aaa aca gtg   438
Cys Gln Ser Asp Asp Pro Ala Ile Ile Leu Gln Glu Gln Lys Thr Val
                1   5   10
cta gcc tct gtt ttt tca gtg ttg tct gcc atc tat gcc tca cag act   486
Leu Ala Ser Val Phe Ser Val Leu Ser Ala Ile Tyr Ala Ser Gln Thr
    15   20   25
gag caa gag tat cta aag ata gaa aaa gta gat ctt cct cta att gac   534
Glu Gln Glu Tyr Leu Lys Ile Glu Lys Val Asp Leu Pro Leu Ile Asp
30   35   40   45
agc ctc att cgg gtc tta caa aat atg gaa cag tgt cag aaa aaa cca   582
Ser Leu Ile Arg Val Leu Gln Asn Met Glu Gln Cys Gln Lys Lys Pro
                50   55   60
gag aac tcg gca gga gtc taacacagag gaaactaaaa ggactgattt   630
Glu Asn Ser Ala Gly Val
                65
aacccaagat gatttccact tgaaaatctt aaaggatat tgttatggtg aagtttctgt   690
ctaataattt ttcaggcatt aacaaaggag acggtggctc agggagtaaa ggaaggccgt   750
tgagcaaaca gaagtgttcc tctgcaattt caaaarcctt cttctttcta tagcccctgt   810
gggtggaaga ttttattaaa atcctacgtg aagttgataa ggcgcttgct kgatgacttg   870
gaaaaaaamc ttcccaagtt tgaaggttca gaastaaaaa rscktgaatg ggaattactt   930
sstgtbcaag aaaataaact ttattttct cactgaaaaa aaaaaaaa   978

<210> 77
<211> 587
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 421..465
<223> Von Heijne matrix
       score 3.90000009536743
       seq LVPLGQSFPLSEP/RC

103

```
<220>
<221> polyA_signal
<222> 553..558
<220>
<221> polyA_site
<222> 575..587
<220>
<221> misc_feature
<222> 182..322
<223> homology
      id :T35951
      est
<220>
<221> misc_feature
<222> 32..132
<223> homology
      id :T35951
      est
<220>
<221> misc_feature
<222> 136..193
<223> homology
      id :T35951
      est
<220>
<221> misc_feature
<222> 182..322
<223> homology
      id :T35949
      est
<220>
<221> misc_feature
<222> 32..132
<223> homology
      id :T35949
      est
<220>
<221> misc_feature
<222> 136..193
<223> homology
      id :T35949
      est
<220>
<221> misc_feature
<222> 136..299
<223> homology
      id :AA381111
      est
<220>
<221> misc_feature
<222> 32..132
<223> homology
      id :AA381111
      est
<220>
<221> misc_feature
<222> 136..322
<223> homology
```

```
          id :AA381001
       est
<220>
<221> misc_feature
<222> 85..132
<223> homology
          id :AA381001
       est
<220>
<221> misc_feature
<222> 182..322
<223> homology
          id :HSCZQE041
       est
<220>
<221> misc_feature
<222> 136..193
<223> homology
          id :HSCZQE041
       est
<220>
<221> misc_feature
<222> 82..132
<223> homology
          id :HSCZQE041
       est
<220>
<221> misc_feature
<222> 316..428
<223> homology
          id :AA477628
       est
<220>
<221> misc_feature
<222> 475..554
<223> homology
          id :AA477628
       est
<220>
<221> misc_feature
<222> 182..322
<223> homology
          id :HSC34G011
       est
<220>
<221> misc_feature
<222> 136..192
<223> homology
          id :HSC34G011
       est
<220>
<221> misc_feature
<222> 41..119
<223> homology
          id :AA090647
       est
<220>
<221> misc_feature
```

<222> 136..184
<223> homology
    id :AA090647
    est
<220>
<221> misc_feature
<222> 316..426
<223> homology
    id :AA505962
    est
<400> 77

```
aattcatttt tcactcctcc ctcctaggtc acacttttca gaaaaagaat ctgcatcctg      60
gaaaccagaa gaaaaatatg agacggggaa tcatcgtgtg atgtgtgtgc tgcctttggc     120
tkwgtgtgtk gaagtycckg ctcaggtgtt aggtacagtg tgtttgatcg tggtggcttg     180
aggggaaccc gctgttcaga gctgtgactg cggctgcact cagagaagct gcccttggct     240
gctcgtagcg ccgggccttc tctcctcgtc atcatccaga gcagccagtg tccgggaggc     300
agaagatgcc ccactccagc ctctggactg ggggctctct tcagtggctg aatgtccagc     360
agagctattt ccttccacag ggggccttgc agggaagggt ccaggacttg acatcttaag     420
```

```
atg cgt ctt gtc ccc ttg ggc cag tca ttt ccc ctc tct gag cct cgg      468
Met Arg Leu Val Pro Leu Gly Gln Ser Phe Pro Leu Ser Glu Pro Arg
-15                     -10                 -5                  1
tgt ctt caa cct gtg aaa tgg gat cat aat cac tgc ctt acc tcc ctc      516
Cys Leu Gln Pro Val Lys Trp Asp His Asn His Cys Leu Thr Ser Leu
              5                 10                  15
acg gtt gtt gtg agg act gag tgt gtg gaa gtt ttt cat aaa ctt tgg      564
Thr Val Val Val Arg Thr Glu Cys Val Glu Val Phe His Lys Leu Trp
          20              25                  30
atg cta gtg taaaaaaaaa aaaa                                          587
Met Leu Val
      35
```

<210> 78
<211> 400
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 198..278
<223> Von Heijne matrix
    score 4.90000009536743
    seq CLLSYIALGAIHA/KI
<220>
<221> polyA_signal
<222> 364..369
<220>
<221> polyA_site
<222> 387..400
<400> 78

```
aactttgcct gggtgtcttg cgttctgcac attccggagg accagcttcc ccatcagaag      60
tctgactcca tggaaaccag atggggcaac ggggtggttc tagtgcagac tgtagctgca     120
gctcctctcc acctctagcc tgctcatttc cagctcagaa attctactaa tggcgttttt     180
```

```
tcttcctgaa aaaggaa atg aac agg gtc cct gct gat tct cca aat atg      230
                   Met Asn Arg Val Pro Ala Asp Ser Pro Asn Met
                   -25                     -20
tgt cta atc tgt tta ctg agt tac ata gca ctt gga gcc atc cat gca      278
Cys Leu Ile Cys Leu Leu Ser Tyr Ile Ala Leu Gly Ala Ile His Ala
    -15                 -10                 -5
aaa atc tgt aga aga gca ttc cag gaa gag gga aga gca aat gca aag      326
```

106

```
Lys Ile Cys Arg Arg Ala Phe Gln Glu Glu Gly Arg Ala Asn Ala Lys
1               5                   10                  15
acg ggc gtg aga gct tgg tgc ata cag cca tgg gcc aaa taaagtttcc        375
Thr Gly Val Arg Ala Trp Cys Ile Gln Pro Trp Ala Lys
            20                  25

ttggaatagc caaaaaaaaa aaaaa                                           400


<210> 79
<211> 1166
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 167..229
<223> Von Heijne matrix
      score 5.59999990463257
      seq LVLSLQFLLLSYD/LF
<220>
<221> polyA_signal
<222> 1133..1138
<220>
<221> polyA_site
<222> 1154..1166
<220>
<221> misc_feature
<222> 22..377
<223> homology
      id :AA306911
      est
<220>
<221> misc_feature
<222> 424..540
<223> homology
      id :AA306911
      est
<220>
<221> misc_feature
<222> 376..424
<223> homology
      id :AA306911
      est
<220>
<221> misc_feature
<222> 4..458
<223> homology
      id :AA417777
      est
<220>
<221> misc_feature
<222> 10..447
<223> homology
      id :AA236327
      est
<220>
<221> misc_feature
<222> 279..714
<223> homology
      id :AA410332
```

```
        est
<220>
<221> misc_feature
<222> 680..893
<223> homology
        id :N32991
        est
<220>
<221> misc_feature
<222> 881..1023
<223> homology
        id :N32991
        est
<220>
<221> misc_feature
<222> 1056..1109
<223> homology
        id :N32991
        est
<220>
<221> misc_feature
<222> 1122..1153
<223> homology
        id :N32991
        est
<220>
<221> misc_feature
<222> 1024..1054
<223> homology
        id :N32991
        est
<220>
<221> misc_feature
<222> 703..893
<223> homology
        id :N24951
        est
<220>
<221> misc_feature
<222> 881..1023
<223> homology
        id :N24951
        est
<220>
<221> misc_feature
<222> 1056..1109
<223> homology
        id :N24951
        est
<220>
<221> misc_feature
<222> 1122..1153
<223> homology
        id :N24951
        est
<220>
<221> misc_feature
<222> 1024..1054
```

237

<223> homology
        id :N24951
        est
<220>
<221> misc_feature
<222> 225..563
<223> homology
        id :AA455215
        est
<220>
<221> misc_feature
<222> 544..631
<223> homology
        id :AA455215
        est
<220>
<221> misc_feature
<222> 629..660
<223> homology
        id :AA455215
        est
<220>
<221> misc_feature
<222> 680..793
<223> homology
        id :N66437
        est
<220>
<221> misc_feature
<222> 1055
<223> n=a, g, c or t
<400> 79

```
aatgacaacc gacgttggag tttggaggtg cttgccttag agcaagggaa acagctctca        60
ttcaaaggaa ctagaagcct ctccctcagt ggtagggaga cagccaggag cggttttctg       120
ggaactgtgg gatgtgccct tggggccccg agaaaacaga aggaag atg ctc cag          175
                                                    Met Leu Gln
                                                        -20

acc agt aac tac agc ctg gtg ctc tct ctg cag ttc ctg ctg ctg tcc        223
Thr Ser Asn Tyr Ser Leu Val Leu Ser Leu Gln Phe Leu Leu Leu Ser
            -15             -10                 -5
tat gac ctc ttt gtc aat tcc ttc tca gaa ctg ctc caa aag act cct        271
Tyr Asp Leu Phe Val Asn Ser Phe Ser Glu Leu Leu Gln Lys Thr Pro
         1               5                 10
gtc atc cag ctt gtg ctc ttc atc atc cag gat att gca gtc ctc ttc        319
Val Ile Gln Leu Val Leu Phe Ile Ile Gln Asp Ile Ala Val Leu Phe
15                  20                 25                  30
aac atc atc atc att ttc ctc atg ttc ttc aac acc tcc gtc ttc cag        367
Asn Ile Ile Ile Ile Phe Leu Met Phe Phe Asn Thr Ser Val Phe Gln
                    35                 40                  45
gct ggc ctg gtc aac ctc cta ttc cat aag ttc aaa ggg acc atc atc        415
Ala Gly Leu Val Asn Leu Leu Phe His Lys Phe Lys Gly Thr Ile Ile
            50                 55                  60
ctg aca gct gtg tac ttt gcc ctc agc atc tcc ctt cat gtc tgg gtc        463
Leu Thr Ala Val Tyr Phe Ala Leu Ser Ile Ser Leu His Val Trp Val
            65                 70                  75
atg aac tta cgc tgg aaa aac tcc aac agc ttc ata tgg aca gat gga        511
Met Asn Leu Arg Trp Lys Asn Ser Asn Ser Phe Ile Trp Thr Asp Gly
         80                 85                  90
```

```
ctt caa atg ctg ttt gta ttc cag aga cta gca gca gtg ttg tac tgc      559
Leu Gln Met Leu Phe Val Phe Gln Arg Leu Ala Ala Val Leu Tyr Cys
95                100                105                110
tac ttc tat aaa cgg aca gcc gta aga cta ggc gat cct cac ttc tac      607
Tyr Phe Tyr Lys Arg Thr Ala Val Arg Leu Gly Asp Pro His Phe Tyr
115                120                125
cag gac tct ttg tgg ctg cgc aag gag ttc atg caa gtt cga agg          652
Gln Asp Ser Leu Trp Leu Arg Lys Glu Phe Met Gln Val Arg Arg
130                135                140
tgacctcttg tcacactgat ggatactttt ccttcctgat agaagccaca tttgctgctt      712
tgcagggaga gttggcccta tgcatgggca aacagctgga ctttccaagg aaggttcaga      772
ctagctgtgt tcagcattca agaaggaaga tcccccctct tgcacaatta gagtgtcccc      832
atcggtctcc agtgcggcat cccttccttg ccttctacct ctgttccacc cccttccttc      892
ctctcctctc tgtaccattc attctccctg accggccttt cttgccgagg gttctgtggc      952
tcttacccttt gtgaagcttt tcctttagcc tgggacagaa ggacctcccg gcccccaaag     1012
gatctcccag wtgaccaaag gatgcgaaga gtgatagtta cgntgctcct gactgatcac     1072
accgcagaca tttagatttt tatacccaag gcactttaaa aaaatgtttt ataaatagag     1132
aataaattga attyttgttc caaaaaaaaa aaaa                                 1166

<210> 80
<211> 754
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 180..383
<223> Von Heijne matrix
      score 4.59999990463257
      seq LPFSLVSMLVTQG/LV
<220>
<221> polyA_signal
<222> 722..727
<220>
<221> polyA_site
<222> 743..754
<220>
<221> misc_feature
<222> 116..450
<223> homology
      id :W68799
      est
<220>
<221> misc_feature
<222> 593..710
<223> homology
      id :W68799
      est
<220>
<221> misc_feature
<222> 18..117
<223> homology
      id :W68799
      est
<220>
<221> misc_feature
<222> 561..598
<223> homology
      id :W68799
```

239

```
        est
<220>
<221> misc_feature
<222> 48..511
<223> homology
        id :AA149518
        est
<220>
<221> misc_feature
<222> 593..673
<223> homology
        id :AA149518
        est
<220>
<221> misc_feature
<222> 535..710.
<223> homology
        id :W80356
        est
<220>
<221> misc_feature
<222> 256..405
<223> homology
        id :W80356
        est
<220>
<221> misc_feature
<222> 432..511
<223> homology
        id :W80356
        est
<220>
<221> misc_feature
<222> 392..437
<223> homology
        . id :W80356
        est
<220>
<221> misc_feature
<222> 535..710
<223> homology
        id :W80631
        est
<220>
<221> misc_feature
<222> 289..437
<223> homology
        id :W80631
        est
<220>
<221> misc_feature
<222> 432..511
<223> homology
        id :W80631
        est
<220>
<221> misc_feature
<222> 343..511
```

```
<223> homology
     id :AA142865
     est
<220>
<221> misc_feature
<222> 535..710
<223> homology
     id :AA142865
     est
<220>
<221> misc_feature
<222> 256..341
<223> homology
     id :AA142865
     est
<220>
<221> misc_feature
<222> 248..511
<223> homology
     id :AA405876
     est
<220>
<221> misc_feature
<222> 21..271
<223> homology
     id :AA405876
     est
<220>
<221> misc_feature
<222> 121..450
<223> homology
     id :W68728
     est
<220>
<221> misc_feature
<222> 592..710
<223> homology
     id :W68728
     est
<220>
<221> misc_feature
<222> 725
<223> n=a, g, c or t
<400> 80
aagacaggtg gggtactcgg gaagctggag cgggccggcg gtgcagtcac gggggagcga       60
ggcctgctgg gcttggcaac gagggactcg gcctcggagg cgacccagac cacacagaca      120
ctgggtcaag gagtaagcag aggataaaca actggaagga gagcaagcac aaagtcatc       179
atg gct tca gcg tct gct cgt gga aac caa gat aaa gat gcc cat ttt        227
Met Ala Ser Ala Ser Ala Arg Gly Asn Gln Asp Lys Asp Ala His Phe
              -65                 -60                 -55
cca cca cca agc aag cag agc ctg ttg ttt tgt cca aaa tca aaa ctg        275
Pro Pro Pro Ser Lys Gln Ser Leu Leu Phe Cys Pro Lys Ser Lys Leu
         -50                 -45                 -40
cac atc cac aga gca gag atc tca aag att atg cga gaa tgt cag gaa        323
His Ile His Arg Ala Glu Ile Ser Lys Ile Met Arg Glu Cys Gln Glu
    -35                 -30                 -25
gaa agt ttc tgg aag aga gct ctg cct ttt tct ctt gta agc atg ctt        371
Glu Ser Phe Trp Lys Arg Ala Leu Pro Phe Ser Leu Val Ser Met Leu
```

```
      -20                      -15                       -10                         -5
gtc acc cag gga cta gtc tac caa ggt tat ttg gca gct aat tct aga        419
Val Thr Gln Gly Leu Val Tyr Gln Gly Tyr Leu Ala Ala Asn Ser Arg
                          1                         5                          10
ttt gga tca ttg ccc aaa gtt gca ctt gct ggt ctc ttg gga ttt ggc        467
Phe Gly Ser Leu Pro Lys Val Ala Leu Ala Gly Leu Leu Gly Phe Gly
              15                        20                        25
ctt gga aag gta tca tac ata gga gta tgc cag agt aaa ttc cat ttt        515
Leu Gly Lys Val Ser Tyr Ile Gly Val Cys Gln Ser Lys Phe His Phe
        30                        35                        40
ttt gaa gat cag ctc cgt ggg gct ggt ttt ggt ccw aca gca             557
Phe Glu Asp Gln Leu Arg Gly Ala Gly Phe Gly Pro Thr Ala
45                        50                        55
taacaggcac tgcctcctta cctgtgagga atgcaaaata aagcatggat taagtgagaa       617
gggagactct cagccttcag cttcctaaat tctgtgtctg tgactttcga agtttttaa        677
acctctgaat ttgtacacat ttaaaatttc aaggtgtact ttaaaatnaa aatacttcta       737
atgtvaaaaa aaaaaaa                                                       754


<210> 81
<211> 709
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 179..298
<223> Von Heijne matrix
      score 4.30000019073486
      seq ITLVSAAPGKVIC/EM
<220>
<221> polyA_signal
<222> 680..685
<220>
<221> polyA_site
<222> 697..708
<220>
<221> misc_feature
<222> 137..291
<223> homology
      id :AA121372
      est
<220>
<221> misc_feature
<222> 6..91
<223> homology
      id :AA121372
      est
<220>
<221> misc_feature
<222> 318..397
<223> homology
      id :AA121372
      est
<220>
<221> misc_feature
<222> 95..132
<223> homology
      id :AA121372
      est
```

```
<220>
<221> misc_feature
<222> 460..501
<223> homology
       id :AA121372
       est
<220>
<221> misc_feature
<222> 432..465
<223> homology
       id :AA121372
       est
<220>
<221> misc_feature
<222> 284..313
<223> homology
       id :AA121372
       est
<220>
<221> misc_feature
<222> 254..670
<223> homology
       id :AA614605
       est
<220>
<221> misc_feature
<222> 392..658
<223> homology
       id :T55234
       est
<220>
<221> misc_feature
<222> 271..327
<223> homology
       id :T55234
       est
<220>
<221> misc_feature
<222> 358..670
<223> homology
       id :AA121362
       est
<220>
<221> misc_feature
<222> 312..344
<223> homology
       id :AA121362
       est
<220>
<221> misc_feature
<222> 2..102
<223> homology
       id :T53974
       est
<220>
<221> misc_feature
<222> 150..258
<223> homology
```

```
        id :T53974
      est
<220>
<221> misc_feature
<222> 95..171
<223> homology
        id :T53974
      est
<220>
<221> misc_feature
<222> 322..628
<223> homology
        id :HSPD02295
      est
<220>
<221> misc_feature
<222> 445..670
<223> homology
        id :AA454502
      est
<220>
<221> misc_feature
<222> 2..102
<223> homology
        id :R09314
      est
<220>
<221> misc_feature
<222> 95..171
<223> homology
        id :R09314
      est
<220>
<221> misc_feature
<222> 150..222
<223> homology
        id :R09314
      est
<400> 81
aaaatcgcgg accaccgggg ctgccakctc gcctgactcc cggcctcttg cgctcctagg    60
ggcggagaag ggtgcgggct cttcgccctt tgtgtccttc tttcactaac ttctggactt   120
tccagctctt ccgaagttcg ttcttgcgca aagcccaaag gctggaaaac cgtccacg     178
atg acc agc atg act cag tct ctg cgg gag gtg ata aag gcc atg acc     226
Met Thr Ser Met Thr Gln Ser Leu Arg Glu Val Ile Lys Ala Met Thr
-40                 -35                 -30                 -25
aag gct cgc aat ttt gag aga gtt ttg gga aag att act ctt gtc tct     274
Lys Ala Arg Asn Phe Glu Arg Val Leu Gly Lys Ile Thr Leu Val Ser
                -20                 -15                 -10
gct gct cct ggg aaa gtg att tgt gaa atg aaa gta gaa gaa gag cat     322
Ala Ala Pro Gly Lys Val Ile Cys Glu Met Lys Val Glu Glu Glu His
            -5              1               5
acc aat gca ata ggc act ctc cac ggc ggt ttg aca gcc acg tta gta     370
Thr Asn Ala Ile Gly Thr Leu His Gly Gly Leu Thr Ala Thr Leu Val
    10              15                  20
gat aac ata tca aca atg gct ctg cta tgc acg gaa agg gga gca ccc     418
Asp Asn Ile Ser Thr Met Ala Leu Leu Cys Thr Glu Arg Gly Ala Pro
25                  30                  35                  40
gga gtc agt gtc gat atg aac ata acg tac atg tca cct gca aaa tta     466
```

115

```
Gly Val Ser Val Asp Met Asn Ile Thr Tyr Met Ser Pro Ala Lys Leu
                45                  50                  55
gga gag gat ata gtg att aca gca cat gtt ctg aag caa gga aaa aca        514
Gly Glu Asp Ile Val Ile Thr Ala His Val Leu Lys Gln Gly Lys Thr
                60                  65                  70
ctt gca ttt acc tct gtg ggt ctg acc aac aag gcc aca gga aaa tta        562
Leu Ala Phe Thr Ser Val Gly Leu Thr Asn Lys Ala Thr Gly Lys Leu
        75                  80                  85
ata gca caa gga aga cac aca aaa cac ctg gga aac tgagagaaca             608
Ile Ala Gln Gly Arg His Thr Lys His Leu Gly Asn
    90                  95                  100
gcagaatgac ctaaagaaac ccaacaatga atatcaagta tagatttgac tcaaacaatt      668
gtaatttttg aaataaacta gcaaaccaa aaaaaaaaaa g                           709


<210> 82
<211> 243
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 100..171
<223> Von Heijne matrix
        score 3.70000004768372
        seq ILFNLLIFLCGFT/NY
<220>
<221> polyA_signal
<222> 211..216
<220>
<221> polyA_site
<222> 230..243
<220>
<221> misc_feature
<222> 2..164
<223> homology
        id :H64488
        est
<220>
<221> misc_feature
<222> 2..164
<223> homology
        id :AA131065
        est
<220>
<221> misc_feature
<222> 5..164
<223> homology
        id :AA224847
        est
<220>
<221> misc_feature
<222> 10..164
<223> homology
        id :AA161042
        est
<220>
<221> misc_feature
<222> 2..84
<223> homology
```

116

```
        id :AA088770
      est
<220>
<221> misc_feature
<222> 104..164
<223> homology
        id :AA088770
      est
<220>
<221> misc_feature
<222> 10..164
<223> homology
        id :AA100852
      est
<220>
<221> misc_feature
<222> 79..164
<223> homology
        id :AA146774
      est
<220>
<221> misc_feature
<222> 79..164
<223> homology
        id :AA146605
      est
<220>
<221> misc_feature
<222> 109..164
<223> homology
        id :AA299239
      est
<220>
<221> misc_feature
<222> 158..207
<223> homology
        id :AA037885
      est
<220>
<221> misc_feature
<222> 160..207
<223> homology
        id :AA480512
      est
<220>
<221> misc_feature
<222> 160..207
<223> homology
        id :AA468030
      est
<220>
<221> misc_feature
<222> 160..207
<223> homology
        id :AA420727
      est
<220>
<221> misc_feature
```

<222> 160..207
<223> homology
       id :AA574382
       est
<220>
<221> misc_feature
<222> 160..207
<223> homology
       id :AA133048
       est
<220>
<221> misc_feature
<222> 200..229
<223> homology
       id :AA469266
       est
<220>
<221> misc_feature
<222> 200..229
<223> homology
       id :AA550735
       est
<220>
<221> misc_feature
<222> 200..229
<223> homology
       id :AA601071
       est
<220>
<221> misc_feature
<222> 200..229
<223> homology
       id :AA225190
       est
<400> 82
aactcagtgg caacacccgg gagctgtttt gtcctttgtg gagcctcagc agttccctct   60
ttcagaactc actgccaaga gccctgaaca ggagccacc atg cag tgc ttc agc   114
                                            Met Gln Cys Phe Ser
                                                              -20
ttc att aag acc atg atg atc ctc ttc aat ttg ctc atc ttt ctg tgt   162
Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu Leu Ile Phe Leu Cys
                -15   -10   -5
ggc ttc acc aac tat acg gat ttt gag gac tca ccc tac ttc aaa atg   210
Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp Ser Pro Tyr Phe Lys Met
                1   5   10
cat aaa cct gtt aca atg taaaaaaaa aaaaa   243
His Lys Pro Val Thr Met
    15

<210> 83
<211> 829
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 346..408
<223> Von Heijne matrix
      score 5.5

118 247

```
        seq SFLPSALVIWTSA/AF
<220>
<221> polyA_signal
<222> 792..797
<220>
<221> polyA_site
<222> 817..829
<220>
<221> misc_feature
<222> 260..464
<223> homology
        id :H57434
        est
<220>
<221> misc_feature
<222> 118..184
<223> homology
        id :H57434
        est
<220>
<221> misc_feature
<222> 56..113
<223> homology
        id :H57434
        est
<220>
<221> misc_feature
<222> 454..485
<223> homology
        id :H57434
        est
<220>
<221> misc_feature
<222> 118..545
<223> homology
        id :N27248
        est
<220>
<221> misc_feature
<222> 65..369
<223> homology
        id :H94779
        est
<220>
<221> misc_feature
<222> 471..519
<223> homology
        id :H94779
        est
<220>
<221> misc_feature
<222> 61..399
<223> homology
        id :H09880
        est
<220>
<221> misc_feature
<222> 408..452
```

```
<223> homology
        id :H09880
      est
<220>
<221> misc_feature
<222> 60..399
<223> homology
        id :H29351
      est
<220>
<221> misc_feature
<222> 393..432
<223> homology
        id :H29351
      est
<220>
<221> misc_feature
<222> 260..444
<223> homology
        id :AA459511
      est
<220>
<221> misc_feature
<222> 449..545
<223> homology
        id :AA459511
      est
<220>
<221> misc_feature
<222> 117..184
<223> homology
        id :AA459511
      est
<220>
<221> misc_feature
<222> 122..399
<223> homology
        id :T74091
      est
<220>
<221> misc_feature
<222> 393..434
<223> homology
        id :T74091
      est
<220>
<221> misc_feature
<222> 61..378
<223> homology
        id :HSC3CB081
      est
<220>
<221> misc_feature
<222> 118..399
<223> homology
        id :T82010
      est
<220>
```
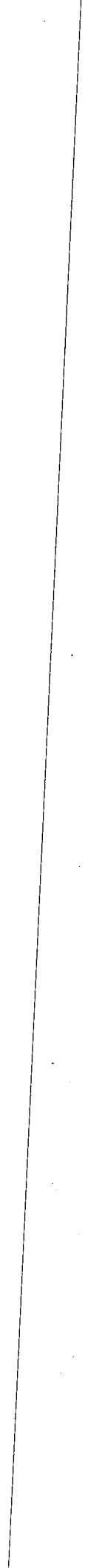
```
<221> misc_feature
<222> 268..545
<223> homology
      id :W02860
      est
<220>
<221> misc_feature
<222> 268..545
<223> homology
      id :N44490
      est
<220>
<221> misc_feature
<222> 115
<223> n=a, g, c or t
<400> 83
actcctttta gcataggggc ttcggcgcca gcggccagcg ctagtcggtc tggtaagtgc    60
ctgatgccga gttccgtctc tcgcgtcttt tcctggtccc aggcaaagcg gasgnagatc   120
ctcaaacggc ctagtgcttc gcgcttccgg agaaaatcag cggtctaatt aattcctctg   180
gtttgttgaa gcagttacca agaatcttca accctttccc acaaaagcta attgagtaca   240
cgttcctgtt gagtacacgt tcctgttgat ttacaaaagg tgcaggtatg agcaggtctg   300
aagactaaca ttttgtgaag ttgtaaaaca gaaaacctgt tagaa atg tgg tgg ttt   357
                                                  Met Trp Trp Phe
                                                      -20
cag caa ggc ctc agt ttc ctt cct tca gcc ctt gta att tgg aca tct   405
Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr Ser
        -15                 -10                 -5
gct gct ttc ata ttt tca tac att act gca gta aca ctc cac cat ata   453
Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His Ile
    1                   5                   10                  15
gac ccg gct tta cct tat atc agt gac act ggt aca gta gct cca gaa   501
Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro Glu
                20                  25                  30
aaa tgc tta ttt ggg gca atg cta aat att gcg gca gtc tta tgt caa   549
Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys Gln
            35                  40                  45
aaa tagaaatcag gaagataatt caacttaaag aagttcattt catgaccaaa         602
Lys
ctcttcagaa acatgtcttt acaagcatat ctcttgtatt gctttctaca ctgttgaatt   662
gtctggcaat atttctgcag tggaaaattt gatttagcta gttcttgact tggataaata   722
tggtaaggtg ggcttttccc cctgtgtaat tggctacsac gtcttacttg agccaagttg   782
gtaagttgaa ataaaatgat watgagagtg acacavaaaa aaaaaaa                 829

<210> 84
<211> 674
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 177..233
<223> Von Heijne matrix
      score 6.09999990463257
      seq LALLWSLPASDLG/RS
<220>
<221> polyA_signal
<222> 644..649
<220>
<221> polyA_site
```

121

```
<222> 663..674
<220>
<221> misc_feature
<222> 194..592
<223> homology
       id :AA496246
       est
<220>
<221> misc_feature
<222> 1..100
<223> homology
       id :AA496246
       est
<220>
<221> misc_feature
<222> 99..202
<223> homology
       id :AA496246
       est
<220>
<221> misc_feature
<222> 187..592
<223> homology
       id :AA476481
       est
<220>
<221> misc_feature
<222> 594..661
<223> homology
       id :AA476481
       est
<220>
<221> misc_feature
<222> 188..592
<223> homology
       id :AA496245
       est
<220>
<221> misc_feature
<222> 594..661
<223> homology
       id :AA496245
       est
<220>
<221> misc_feature
<222> 194..444
<223> homology
       id :AA476480
       est
<220>
<221> misc_feature
<222> 1..102
<223> homology
       id :AA476480
       est
<220>
<221> misc_feature
<222> 99..187
```

<223> homology
        id :AA476480
        est
<220>
<221> misc_feature
<222> 437..592
<223> homology
        id :AA505488
        est
<220>
<221> misc_feature
<222> 594..661
<223> homology
        id :AA505488
        est
<220>
<221> misc_feature
<222> 441..592
<223> homology
        id :AA554685
        est
<220>
<221> misc_feature
<222> 594..661
<223> homology
        id :AA554685
        est
<220>
<221> misc_feature
<222> 414..503
<223> homology
        id :AA215595
        est
<220>
<221> misc_feature
<222> 510..539
<223> homology
        id :AA215595
        est
<400> 84

```
ataagtgaac cagaccaccc tgatggcatc cacagtgatg tcaaggttgg ggctggccag      60
gggtgggtgg actagaagca tttgggagta gtggccaggg gccctggacg ctagccacgg     120
agctgctgca cagagcctgg tgtccacaag cttccaggtt ggggttggag cctggg atg     179
                                                                  Met
agc ccc ggc agc gcc ttg gcc ctt ctg tgg tcc ctg cca gcc tct gac      227
Ser Pro Gly Ser Ala Leu Ala Leu Leu Trp Ser Leu Pro Ala Ser Asp
            -15                 -10                 -5
ctg ggc cgg tca gtc att gct gga ctc tgg cca cac act ggc gtt ctc      275
Leu Gly Arg Ser Val Ile Ala Gly Leu Trp Pro His Thr Gly Val Leu
         1              5               10
atc cac ttg gaa aca agc cag tct ttt ctg caa ggt cag ttg acc aag      323
Ile His Leu Glu Thr Ser Gln Ser Phe Leu Gln Gly Gln Leu Thr Lys
15              20              25                      30
agc ata ttt ccc ctc tgt tgt aca tcg ttg ttt tgt gtt tgt gtt gta      371
Ser Ile Phe Pro Leu Cys Cys Thr Ser Leu Phe Cys Val Cys Val Val
             35                 40                  45
aca gtg ggt gga ggg agg gtg ggg tct aca ttt gtt gca tgagtcgatg      420
Thr Val Gly Gly Gly Arg Val Gly Ser Thr Phe Val Ala
```

123

```
                    50                      55
ggtcagaact ttagtatacg catgcgtcct ctgagtgaca gggcattttg tcgaaaataa      480
gcaccttggt aactaaaccc ctctaatagc tataaaggct ttagttctgt attgattaag      540
ttactgtaaa agcttgggtt tatttttgta ggacttaatg gctaagaatt agggaacata      600
gcaagggggc tcctctgttg gagtaatgta aattgtaatt ataaataaac atgcaaacct      660
ttaaaaaaaa aaaa                                                         674


<210> 85
<211> 478
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 179..319
<223> Von Heijne matrix
      score 5.5
      seq SALLFFARPCVFC/FK
<220>
<221> polyA_signal
<222> 461..466
<220>
<221> polyA_site
<222> 465..478
<220>
<221> misc_feature
<222> 2..464
<223> homology
       id :AA310996
      est
<220>
<221> misc_feature
<222> 8..464
<223> homology
       id :AA312901
      est
<220>
<221> misc_feature
<222> 2..416
<223> homology
       id :AA401411
      est
<220>
<221> misc_feature
<222> 2..349
<223> homology
       id :R64030
      est
<220>
<221> misc_feature
<222> 56..464
<223> homology
       id :AA400108
      est
<220>
<221> misc_feature
<222> 126..273
<223> homology
       id :AA010825
```

```
        est
<220>
<221> misc_feature
<222> 2..147
<223> homology
      id :AA010825
        est
<220>
<221> misc_feature
<222> 358..435
<223> homology
      id :AA010825
        est
<220>
<221> misc_feature
<222> 78..464
<223> homology
      id :AA504732
        est
<220>
<221> misc_feature
<222> 90..441
<223> homology
      id :H60506
        est
<220>
<221> misc_feature
<222> 59..349
<223> homology
      id :AA346780
        est
<220>
<221> misc_feature
<222> 2..331
<223> homology
      id :AA281167
        est
<220>
<221> misc_feature
<222> 6..236
<223> homology
      id :R35805
        est
<220>
<221> misc_feature
<222> 232..284
<223> homology
      id :R35805
        est
<220>
<221> misc_feature
<222> 41..307
<223> homology
      id :H13784
        est
<220>
<221> misc_feature
<222> 2..40
```

<223> homology
      id :H13784
      est
<220>
<221> misc_feature
<222> 64..280
<223> homology
      id :AA128122
      est
<220>
<221> misc_feature
<222> 293..349
<223> homology
      id :AA128122
      est
<220>
<221> misc_feature
<222> 332..385
<223> homology
      id :AA128122
      est
<220>
<221> misc_feature
<222> 163..420
<223> homology
      id :AA555127
      est
<400> 85

```
aagtccttcg cgccctcctc gccctcccca ccgacatcat gctccagttc ctgcttggat      60
ttacactggg caacgtggtt ggaatgtatc tggctcagaa ctatgatata ccaaacctgg     120
ctaaaaaact tgaagaaatt aaaaaggact tggatgccaa gaagaaaccc cctagtgc       178
atg aga ctg cct cca gca ctg cct tca gga tat act gat tct act gct      226
Met Arg Leu Pro Pro Ala Leu Pro Ser Gly Tyr Thr Asp Ser Thr Ala
        -45                   -40                   -35
ctt gag ggc ctc gtt tac tat ctg aac caa aag ctt ttg ttt tcg tct      274
Leu Glu Gly Leu Val Tyr Tyr Leu Asn Gln Lys Leu Leu Phe Ser Ser
        -30                   -25                   -20
cca gcc tca gca ctt ctc ttc ttt gct aga ccc tgt gtt ttt tgc ttt      322
Pro Ala Ser Ala Leu Leu Phe Phe Ala Arg Pro Cys Val Phe Cys Phe
-15                   -10                   -5                    1
aaa gca agc aaa atg ggg ccc caa ttt gag aac tac cca aca ttt cca      370
Lys Ala Ser Lys Met Gly Pro Gln Phe Glu Asn Tyr Pro Thr Phe Pro
            5                   10                    15
aca tac tca cct ctt ccc ata atc cct ttc caa ctg cat ggg agg ttc      418
Thr Tyr Ser Pro Leu Pro Ile Ile Pro Phe Gln Leu His Gly Arg Phe
        20                    25                    30
taagactgga attatggtgc tagattagta aacatgactt ttaatgaaaa aaaaacaaaa     478
```

<210> 86
<211> 952
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 112..237
<223> Von Heijne matrix
      score 7.19999980926514
      seq ILFSLSFLLVIIT/FP

<220>
<221> polyA_signal
<222> 910..915
<220>
<221> polyA_site
<222> 940..952
<400> 86
aatactttct cctctcccct ctcccaagca catctgagtt gctgcctgtt cttcacactt   60
agctccaaac ccatgaaaaa ttgccaagta taaaagcttc tcaagaatga g atg gat   117
                                                          Met Asp
tct agg gtg tct tca cct gag aag caa gat aaa gag aat ttc gtg ggt   165
Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys Glu Asn Phe Val Gly
-40   -35   -30   -25
gtc aac aat aaa cgg ctt ggt gta tgt ggc tgg atc ctg ttt tcc ctc   213
Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp Ile Leu Phe Ser Leu
              -20   -15   -10
tct ttc ctg ttg gtg atc att acc ttc ccc atc tcc ata tgg atg tgc   261
Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile Ser Ile Trp Met Cys
          -5   1   5
ttg aag att tgatcctggt cctgccatgc ataratgtgt ttgtcaaagt   310
Leu Lys Ile
        10
tgacctccga acagttactt gcaacattcc tccacaagag atcctcacca rgagactccg   370
taactactca ggtagatgga gttgtctatt acagaatcta tagtgctgtc tcagcagtgg   430
ctaakgtcaa cgatgtccat caagcaacat ttctgctggc tcaaaccact ctgagaaatg   490
tcktagggac acaggacctt gtccccagat cttaggctgg acgagaagag atcgcccata   550
agcatccaga ctktacttga tgatgccacc gaactggtgg gggatccggg tggcccgagt   610
ggaaatcaaa gatgttcgga ttcccgtgca gttgcagaga tccatggcag ccgaggstga   670
ggccacccgg aaggsgagag ccaaggtcct tgcagctgaa ggagaaatga atgsttccaa   730
atccctgaag tcagcctcca tggtgstggs tgagtytccc atagctytcc agstgsgsta   790
cctgcagacc ttgagcacgg tagccaccga gaagaatttt acgattgtgt ttcctbtgcc   850
catgaatata ctagagggca ttggtggcgt cagstatgat aaccacaaga agsttbscaa   910
ataaagcctg aggtcybctt gcggtagtca aaaaaaaaaa aa   952

<210> 87
<211> 131
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -13..-1
<400> 87
Met Leu Ala Val Ser Leu Thr Val Pro Leu Leu Gly Ala Met Met Leu
            -10   -5   1
Leu Glu Ser Pro Ile Asp Pro Gln Pro Leu Ser Phe Lys Glu Pro Pro
    5   10   15
Leu Leu Leu Gly Val Leu His Pro Asn Thr Lys Leu Arg Gln Ala Glu
20   25   30   35
Arg Leu Phe Glu Asn Gln Leu Val Gly Pro Glu Ser Ile Ala His Ile
            40   45   50
Gly Asp Val Met Phe Thr Gly Thr Ala Asp Gly Arg Val Val Lys Leu
            55   60   65
Glu Asn Gly Glu Ile Glu Thr Ile Ala Arg Phe Gly Ser Gly Pro Cys
        70   75   80
Lys Thr Arg Gly Asp Glu Pro Val Cys Gly Arg Pro Leu Gly Ile Arg
    85   90   95
Gly Arg Ala Gln Trp Asp Ser Leu Cys Gly Arg Cys Ile Gln Arg Asp
100   105   110   115

Tyr Leu Lys

<210> 88
<211> 63
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -35..-1
<400> 88
Met Leu Thr Val Asn Asp Val Arg Phe Tyr Arg Asn Val Arg Ser Asn
-35   -30   -25   -20
His Phe Pro Phe Val Arg Leu Cys Gly Leu Leu His Leu Trp Leu Lys
            -15   -10   -5
Val Phe Ser Leu Lys Gln Leu Lys Lys Lys Ser Trp Ser Lys Tyr Leu
            1   5   10
Phe Glu Ser Cys Cys Tyr Arg Ser Leu Tyr Val Cys Val Phe Ile
        15   20   25

<210> 89
<211> 163
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -31..-1
<220>
<221> UNSURE
<222> 91,108,109,112,124
<223> Xaa = any one of the twenty amino acids
<400> 89
Met Ser Pro Ala Phe Arg Ala Met Asp Val Glu Pro Arg Ala Lys Gly
    -30   -25   -20
Ser Phe Trp Ser Pro Leu Ser Thr Arg Ser Gly Gly Thr His Ala Cys
-15   -10   -5   1
Ser Ala Ser Met Arg Gln Pro Trp Ala Ser Pro Trp Ser Gln Gly Asn
            5   10   15
Ile Ser Ser Thr Arg Pro Ser Leu Leu Arg Cys Ala Asn Ser Leu Pro
            20   25   30
Ser Thr Lys Asp Lys Ala Lys Gly Pro Leu Leu Ala Gly His Pro Cys
        35   40   45
Pro Ile Phe Ser Pro Gly Pro Phe Pro Cys Gly His Arg Glu Val Trp
50   55   60   65
Pro Glu Tyr Pro Thr Pro Ala Pro Leu His Pro Glu Leu Gly Ala Thr
                70   75   80
Ser Glu Val Ser Ser Leu Ser Glu His Xaa Phe Pro Cys Ser Ser Arg
                85   90   95
Gly Leu Ser Arg Leu Ser Asp Ala Gly Ala Xaa Xaa Pro Glu Xaa Lys
        100   105   110
Gly Val Gln Pro Val Val Cys Lys Ala Leu Xaa Gly Thr Ala Glu Thr
        115   120   125
Pro Pro Pro
130

<210> 90
<211> 52
<212> PRT
<213> Homo sapiens

128

<220>
<221> SIGNAL
<222> -32..-1
<400> 90
Met Leu Gly Thr Thr Gly Leu Gly Thr Gln Gly Pro Ser Gln Gln Ala
        -30   -25   -20
Leu Gly Phe Phe Ser Phe Met Leu Leu Gly Met Gly Gly Cys Leu Pro
        -15   -10   -5
Gly Phe Leu Leu Gln Pro Pro Asn Arg Ser Pro Thr Leu Pro Ala Ser
1   5   10   15
Thr Phe Ala His
            20

<210> 91
<211> 124
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -97..-1
<400> 91
Met Ala Asp Asp Leu Lys Arg Phe Leu Tyr Lys Lys Leu Pro Ser Val
        -95   -90   -85
Glu Gly Leu His Ala Ile Val Val Ser Asp Arg Asp Gly Val Pro Val
        -80   -75   -70
Ile Lys Val Ala Asn Asp Asn Ala Pro Glu His Ala Leu Arg Pro Gly
-65   -60   -55   -50
Phe Leu Ser Thr Phe Ala Leu Ala Thr Asp Gln Gly Ser Lys Leu Gly
                -45   -40   -35
Leu Ser Lys Asn Lys Ser Ile Ile Cys Tyr Tyr Asn Thr Tyr Gln Val
                -30   -25   -20
Val Gln Phe Asn Arg Leu Pro Leu Val Val Ser Phe Ile Ala Ser Ser
        -15   -10   -5
Ser Ala Asn Thr Gly Leu Ile Val Ser Leu Glu Lys Glu Leu Ala Pro
    1   5   10   15
Leu Phe Glu Glu Leu Arg Gln Val Val Glu Ile Ser
                20   25

<210> 92
<211> 230
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -24..-1
<220>
<221> UNSURE
<222> 54,79
<223> Xaa = any one of the twenty amino acids
<400> 92
Met Ala Ser Leu Gly Leu Gln Leu Val Gly Tyr Ile Leu Gly Leu Leu
                -20   -15   -10
Gly Leu Leu Gly Thr Leu Val Ala Met Leu Leu Pro Ser Trp Lys Thr
            -5   1   5
Ser Ser Tyr Val Gly Ala Ser Ile Val Thr Ala Val Gly Phe Ser Lys
    10   15   20
Gly Leu Trp Met Glu Cys Ala Thr His Ser Thr Gly Ile Thr Gln Cys
25   30   35   40

129

```
Asp Ile Tyr Ser Thr Leu Leu Gly Leu Pro Ala Asp Ile Xaa Ala Ala
            45                  50              55
Gln Ala Met Met Val Thr Ser Ser Ala Ile Ser Ser Leu Ala Cys Ile
            60                  65              70
Ile Ser Val Val Gly Met Xaa Cys Thr Val Phe Cys Gln Glu Ser Arg
            75                  80              85
Ala Lys Asp Arg Val Ala Val Ala Gly Gly Val Phe Phe Ile Leu Gly
        90                  95                  100
Gly Leu Leu Gly Phe Ile Pro Val Ala Trp Asn Leu His Gly Ile Leu
105                     110                 115                 120
Arg Asp Phe Tyr Ser Pro Leu Val Pro Asp Ser Met Lys Phe Glu Ile
                125                 130                 135
Gly Glu Ala Leu Tyr Leu Gly Ile Ile Ser Ser Leu Phe Ser Leu Ile
            140                 145                 150
Ala Gly Ile Ile Leu Cys Phe Ser Cys Ser Ser Gln Arg Asn Arg Ser
            155                 160                 165
Asn Tyr Tyr Asp Ala Tyr Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser
        170                 175                 180
Pro Arg Pro Gly Gln Pro Pro Lys Val Lys Ser Glu Phe Asn Ser Tyr
185                     190                 195                 200
Ser Leu Thr Gly Tyr Val
                205


<210> 93
<211> 72
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -32..-1
<400> 93
Met Phe Ala Pro Ala Val Met Arg Ala Phe Arg Lys Asn Lys Thr Leu
            -30                 -25                 -20
Gly Tyr Gly Val Pro Met Leu Leu Leu Ile Val Gly Gly Ser Phe Gly
        -15                 -10                 -5
Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp Ala Val Lys Ser Lys Met
1               5                   10                  15
Asp Pro Glu Leu Glu Lys Lys Pro Lys Glu Asn Lys Ile Ser Leu Glu
            20                  25                  30
Ser Glu Tyr Glu Gly Ser Ile Cys
        35                  40


<210> 94
<211> 91
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -36..-1
<400> 94
Met Asn Thr Phe Glu Pro Asp Ser Leu Ala Val Ile Ala Phe Phe Leu
        -35                 -30                 -25
Pro Ile Trp Thr Phe Ser Ala Leu Thr Phe Leu Phe Leu His Leu Pro
-20                 -15                 -10                 -5
Pro Ser Thr Ser Leu Phe Ile Asn Leu Ala Arg Gly Gln Ile Lys Gly
                1               5                   10
Pro Leu Gly Leu Ile Leu Leu Leu Ser Phe Cys Gly Gly Tyr Thr Lys
        15                  20                  25
```

```
Cys Asp Phe Ala Leu Ser Tyr Leu Glu Ile Pro Asn Arg Ile Glu Phe
    30                  35              40
Ser Ile Met Asp Pro Lys Arg Lys Thr Lys Cys
45                  50              55


<210> 95
<211> 106
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -32..-1
<400> 95
Met Phe Ala Pro Ala Val Met Arg Ala Phe Arg Lys Asn Lys Thr Leu
        -30              -25              -20
Gly Tyr Gly Val Pro Met Leu Leu Leu Ile Val Gly Gly Ser Phe Gly
        -15              -10              -5
Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp Ala Val Lys Gly Lys Met
1               5               10              15
Asp Pro Glu Leu Glu Lys Lys Leu Lys Glu Asn Lys Ile Ser Leu Glu
        20              25              30
Ser Glu Tyr Glu Lys Ile Lys Asp Ser Lys Phe Asp Asp Trp Lys Asn
        35              40              45
Ile Arg Gly Pro Arg Pro Trp Glu Asp Pro Asp Leu Leu Gln Gly Arg
        50              55              60
Asn Pro Glu Ser Leu Lys Thr Lys Thr Thr
65                  70


<210> 96
<211> 172
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -21..-1
<400> 96
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
        -20              -15              -10
Ile Trp Thr Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr
-5               1               5               10
Leu His His Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr
            15              20              25
Val Ala Pro Glu Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala
        30              35              40
Val Leu Cys Ile Ala Thr Ile Tyr Val Arg Tyr Lys Gln Val His Ala
        45              50              55
Leu Ser Pro Glu Glu Asn Val Ile Ile Lys Leu Asn Lys Ala Gly Leu
60              65              70              75
Val Leu Gly Ile Leu Ser Cys Leu Gly Leu Ser Ile Val Ala Asn Phe
            80              85              90
Gln Glu Asn Asn Pro Phe Cys Cys Thr Cys Lys Trp Ser Cys Ala Tyr
            95              100             105
Leu Trp Tyr Gly Leu Ile Ile Tyr Val Cys Ser Asp His Pro Phe Leu
        110             115             120
Pro Lys Cys Ser Pro Lys Ser Asn Gly Lys Thr Ser Leu Leu Asp Gln
        125             130             135
Thr Val Val Gly Tyr Leu Val Trp Ser Lys Cys Thr
140             145             150
```

131

<210> 97
<211> 56
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 97

```
Met Cys Phe Pro Glu His Arg Arg Gln Met Tyr Ile Gln Asp Arg Leu
            -40                 -35                 -30
Asp Ser Val Thr Arg Arg Ala Arg Gln Gly Arg Ile Cys Ala Ile Leu
        -25                 -20                 -15
Leu Leu Gln Ser Gln Cys Ala Tyr Trp Ala Leu Pro Glu Pro Arg Thr
-10                     -5                  1                   5
Leu Asp Gly Gly His Leu Met Gln
                10
```

<210> 98
<211> 46
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<400> 98

```
Met Gln Asn His Leu Gln Thr Arg Pro Leu Phe Leu Thr Cys Leu Phe
            -20                 -15                 -10
Trp Pro Leu Ala Ala Leu Asn Val Asn Ser Thr Phe Glu Cys Leu Ile
        -5                  1                   5                   10
Leu Gln Cys Ser Val Phe Ser Phe Ala Phe Phe Ala Leu Trp
                    15                  20
```

<210> 99
<211> 251
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -28..-1
<220>
<221> UNSURE
<222> 54,131,132,140,179,194,213,221
<223> Xaa = any one of the twenty amino acids
<400> 99

```
Met Trp Arg Leu Leu Ala Arg Ala Ser Ala Pro Leu Leu Arg Val Pro
            -25                 -20                 -15
Leu Ser Asp Ser Trp Ala Leu Leu Pro Ala Ser Ala Gly Val Lys Thr
            -10                 -5                  1
Leu Leu Pro Val Pro Ser Phe Glu Asp Val Ser Ile Pro Glu Lys Pro
5                       10                  15                  20
Lys Leu Arg Phe Ile Glu Arg Ala Pro Leu Val Pro Lys Val Arg Arg
                25                  30                  35
Glu Pro Lys Asn Leu Ser Asp Ile Arg Gly Pro Ser Thr Glu Ala Thr
                40                  45                  50
Glu Xaa Thr Glu Gly Asn Phe Ala Ile Leu Ala Leu Gly Gly Gly Tyr
        55                  60                  65
Leu His Trp Gly His Phe Glu Met Met Arg Leu Thr Ile Asn Arg Ser
```

```
        70                      75                      80
Met Asp Pro Lys Asn Met Phe Ala Ile Trp Arg Val Pro Ala Pro Phe
85                      90                      95                      100
Lys Pro Ile Thr Arg Lys Ser Val Gly His Arg Met Gly Gly Gly Lys
                    105                     110                     115
Gly Ala Ile Asp His Tyr Val Thr Pro Val Lys Ala Gly Arg Xaa Xaa
                    120                     125                     130
Val Glu Met Gly Gly Arg Cys Xaa Phe Glu Glu Val Gln Gly Phe Leu
                    135                     140                     145
Asp Gln Val Ala His Lys Leu Pro Phe Ala Ala Lys Ala Val Ser Arg
                    150                     155                     160
Gly Thr Leu Glu Lys Met Arg Lys Asp Gln Glu Glu Arg Glu Xaa Asn
165                     170                     175                     180
Asn Gln Asn Pro Trp Thr Phe Glu Arg Ile Ala Thr Ala Xaa Met Leu
                    185                     190                     195
Gly Ile Arg Lys Val Leu Ser Pro Tyr Asp Leu Thr His Lys Gly Lys
                    200                     205                     210
Xaa Trp Gly Lys Phe Tyr Met Pro Xaa Arg Val
                    215                     220

<210> 100
<211> 77
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -30..-1
<400> 100
Met Leu Arg Leu Asp Ile Ile Asn Ser Leu Val Thr Thr Val Phe Met
-30                     -25                     -20                     -15
Leu Ile Val Ser Val Leu Ala Leu Ile Pro Glu Thr Thr Thr Leu Thr
                    -10                     -5                      1
Val Gly Gly Gly Val Phe Ala Leu Val Thr Ala Val Cys Cys Leu Ala
                5                       10                      15
Asp Gly Ala Leu Ile Tyr Arg Lys Leu Leu Phe Asn Pro Ser Gly Pro
            20                      25                      30
Tyr Gln Lys Lys Pro Val His Glu Lys Lys Glu Val Leu
35                      40                      45

<210> 101
<211> 81
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -31..-1
<400> 101
Met Ser Asn Thr His Thr Val Leu Val Ser Leu Pro His Pro His Pro
        -30                     -25                     -20
Ala Leu Thr Cys Cys His Leu Gly Leu Pro His Pro Val Arg Ala Pro
-15                     -10                     -5                      1
Arg Pro Leu Pro Arg Val Glu Pro Trp Asp Pro Arg Trp Gln Asp Ser
                5                       10                      15
Glu Leu Arg Tyr Pro Gln Ala Met Asn Ser Phe Leu Asn Glu Arg Ser
            20                      25                      30
Ser Pro Cys Arg Thr Leu Arg Gln Glu Ala Ser Ala Asp Arg Cys Asp
    35                      40                      45
Leu
```

133

<210> 102
<211> 126
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -20..-1
<400> 102

```
Met Lys Val His Met His Thr Lys Phe Cys Leu Ile Cys Leu Leu Thr
-20                 -15                 -10                     -5
Phe Ile Phe His His Cys Asn His Cys His Glu Glu His Asp His Gly
                1                 5                     10
Pro Glu Ala Leu His Arg Gln His Arg Gly Met Thr Glu Leu Glu Pro
        15                  20                  25
Ser Lys Phe Ser Lys Gln Ala Ala Glu Asn Glu Lys Lys Tyr Tyr Ile
    30                  35                  40
Glu Lys Leu Phe Glu Arg Tyr Gly Glu Asn Gly Arg Leu Ser Phe Phe
45                  50                  55                      60
Gly Leu Glu Lys Leu Leu Thr Asn Leu Gly Leu Gly Glu Arg Lys Val
            65                  70                  75
Val Glu Ile Asn His Glu Asp Leu Gly His Asp His Val Ser His Leu
        80                  85                  90
Arg Tyr Phe Gly Ser Ser Arg Gly Lys Ala Phe Ser Leu Thr
        95                  100                 105
```

<210> 103
<211> 273
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -45..-1
<220>
<221> UNSURE
<222> 181,187,193,196,198,199,203,212,214
<223> Xaa = any one of the twenty amino acids
<400> 103

```
Met Asn Trp Ser Ile Phe Glu Gly Leu Leu Ser Gly Val Asn Lys Tyr
-45                 -40                 -35                     -30
Ser Thr Ala Phe Gly Arg Ile Trp Leu Ser Leu Val Phe Ile Phe Arg
            -25                 -20                 -15
Val Leu Val Tyr Leu Val Thr Ala Glu Arg Val Trp Ser Asp Asp His
            -10                 -5                  1
Lys Asp Phe Asp Cys Asn Thr Arg Gln Pro Gly Cys Ser Asn Val Cys
    5                   10                  15
Phe Asp Glu Phe Phe Pro Val Ser His Val Arg Leu Trp Ala Leu Gln
20                  25                  30                      35
Leu Ile Leu Val Thr Cys Pro Ser Leu Leu Val Val Met His Val Ala
            40                  45                  50
Tyr Arg Glu Val Gln Glu Lys Arg His Arg Glu Ala His Gly Glu Asn
            55                  60                  65
Ser Gly Arg Leu Tyr Leu Asn Pro Gly Lys Lys Arg Gly Gly Leu Trp
            70                  75                  80
Trp Thr Tyr Val Cys Ser Leu Val Phe Lys Ala Ser Val Asp Ile Ala
            85                  90                  95
Phe Leu Tyr Val Phe His Ser Phe Tyr Pro Lys Tyr Ile Leu Pro Pro
```

```
              100                 105                 110                 115
Val Val Lys Cys His Ala Asp Pro Cys Pro Asn Ile Val Asp Cys Phe
                  120                 125                 130
Ile Ser Lys Pro Ser Glu Lys Asn Ile Phe Thr Leu Phe Met Val Ala
              135                 140                 145
Thr Ala Ala Ile Cys Ile Leu Leu Asn Leu Val Glu Leu Ile Tyr Leu
              150                 155                 160
Val Ser Lys Arg Cys His Glu Cys Leu Ala Ala Arg Lys Ala Gln Ala
          165                 170                 175
Met Xaa Thr Gly His His Pro Xaa Asp Thr Thr Phe Ser Xaa Lys Gln
180                 185                 190                 195
Xaa Asp Xaa Xaa Ser Gly Asp Xaa Ile Phe Leu Gly Ser Asp Ser His
              200                 205                 210
Xaa Pro Xaa Leu Pro Asp Arg Pro Arg Asp His Val Lys Lys Thr Ile
          215                 220                 225
Leu

<210> 104
<211> 158
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -37..-1
<400> 104
Met Ala Ser Lys Ile Leu Leu Asn Val Gln Glu Glu Val Thr Cys Pro
              -35                 -30                 -25
Ile Cys Leu Glu Leu Leu Thr Glu Pro Leu Ser Leu Asp Cys Gly His
          -20                 -15                 -10
Ser Leu Cys Arg Ala Cys Ile Thr Val Ser Asn Lys Glu Ala Val Thr
-5                  1                 5                   10
Ser Met Gly Gly Lys Ser Ser Cys Pro Val Cys Gly Ile Ser Tyr Ser
              15                  20                  25
Phe Glu His Leu Gln Ala Asn Gln His Arg Ala Asn Ile Val Glu Arg
          30                  35                  40
Leu Lys Glu Val Lys Leu Ser Pro Asp Asn Gly Lys Lys Arg Asp Leu
          45                  50                  55
Cys Asp His His Gly Glu Lys Leu Leu Leu Phe Cys Lys Glu Asp Arg
60                  65                  70                  75
Lys Val Ile Cys Trp Leu Cys Glu Arg Ser Gln Glu His Arg Gly His
              80                  85                  90
His Thr Gly Pro His Gly Gly Ser Ile Gln Gly Met Ser Gly Glu Thr
          95                  100                 105
Pro Gly Ser Pro Gln Glu Ala Glu Glu Gly Arg Gly Gly Ser
          110                 115                 120

<210> 105
<211> 51
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -19..-1
<220>
<221> UNSURE
<222> 8
<223> Xaa = any one of the twenty amino acids
<400> 105
```

```
Met Arg Thr Leu Phe Asn Leu Leu Trp Leu Ala Leu Ala Cys Ser Pro
                 -15              -10                    -5
Val His Thr Thr Leu Ser Lys Ser Asp Ala Xaa Lys Pro Pro Gln Arg
             1               5                  10
Arg Cys Trp Arg Arg Val Ser Phe Gln Ile Ser Arg Cys Lys Thr Gly
    15              20                  25
Val Trp Trp
30


<210> 106
<211> 359
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -34..-1
<220>
<221> UNSURE
<222> 20,64,65,130,156,282,288,289,294,296,300,302,310
<223> Xaa = any one of the twenty amino acids
<400> 106
Met Leu Leu Ser Ile Gly Met Leu Met Leu Ser Ala Thr Gln Val Tyr
                 -30              -25                    -20
Thr Ile Leu Thr Val Gln Leu Phe Ala Phe Leu Asn Leu Leu Pro Val
             -15              -10                    -5
Glu Ala Asp Ile Leu Ala Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe
             1               5                  10
Asp Asp Leu Pro Ala Xaa Phe Gly Tyr Arg Leu Pro Ala Glu Gly Leu
15                  20                  25                      30
Lys Gly Phe Leu Ile Asn Ser Lys Pro Glu Asn Ala Cys Glu Pro Ile
                 35                  40                  45
Val Pro Pro Pro Val Lys Asp Asn Ser Ser Gly Thr Phe Ile Val Leu
             50                  55                  60
Ile Xaa Xaa Leu Asp Cys Asn Phe Asp Ile Lys Val Leu Asn Ala Gln
        65                  70                  75
Arg Ala Gly Tyr Lys Ala Ala Ile Val His Asn Val Asp Ser Asp Asp
    80                  85                  90
Leu Ile Ser Met Gly Ser Asn Asp Ile Glu Val Leu Lys Lys Ile Asp
95                  100                 105                     110
Ile Pro Ser Val Phe Ile Gly Glu Ser Ser Ala Ser Ser Leu Lys Asp
                 115                 120                 125
Glu Phe Thr Xaa Glu Lys Gly Gly His Leu Ile Leu Val Pro Glu Phe
             130                 135                 140
Ser Leu Pro Leu Glu Tyr Tyr Leu Ile Pro Phe Leu Ile Xaa Val Gly
         145                 150                 155
Ile Cys Leu Ile Leu Ile Val Ile Phe Met Ile Thr Lys Leu Ser Arg
    160                 165                 170
Asp Arg His Arg Ala Arg Arg Asn Arg Leu Arg Lys Asp Gln Leu Lys
175                 180                 185                     190
Lys Leu Pro Val His Lys Phe Lys Lys Gly Asp Glu Tyr Asp Val Cys
             195                 200                 205
Ala Ile Cys Leu Asp Glu Tyr Glu Asp Gly Asp Lys Leu Arg Ile Leu
             210                 215                 220
Pro Cys Ser His Ala Tyr His Cys Lys Cys Val Asp Pro Trp Leu Thr
         225                 230                 235
Lys Thr Lys Lys Thr Cys Pro Val Cys Arg Gln Lys Val Val Pro Ser
    240                 245                 250
Gln Gly Asp Ser Asp Ser Asp Thr Asp Ser Ser Gln Glu Glu Asn Glu
```

```
255                 260                 265                 270
Val Thr Glu His Thr Pro Leu Leu Arg Pro Leu Xaa Phe Cys Gln Cys
            275                 280                 285
Pro Xaa Xaa Phe Gly Ala Leu Xaa Gly Xaa Pro Ala His Xaa Gln Xaa
            290                 295                 300
His Asp Arg Ile Ile Gln Thr Xaa Glu Glu Asp Asp Asn Glu Asp Thr
            305                 310                 315
Asp Ser Ser Asp Ala Glu Glu
            320                 325
```

```
<210> 107
<211> 291
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 107
```

```
Met Asp Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys Glu Asn Phe
            -40                 -35                 -30
Val Gly Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp Ile Leu Phe
            -25                 -20                 -15
Ser Leu Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile Ser Ile Trp
-10                 -5                  1                   5
Met Cys Leu Lys Ile Ile Lys Glu Tyr Glu Arg Ala Val Val Phe Arg
            10                  15                  20
Leu Gly Arg Ile Gln Ala Asp Lys Ala Lys Gly Pro Gly Leu Ile Leu
            25                  30                  35
Val Leu Pro Cys Ile Asp Val Phe Val Lys Val Asp Leu Arg Thr Val
            40                  45                  50
Thr Cys Asn Ile Pro Pro Gln Glu Ile Leu Thr Arg Asp Ser Val Thr
55                  60                  65                  70
Thr Gln Val Asp Gly Val Val Tyr Tyr Arg Ile Tyr Ser Ala Val Ser
            75                  80                  85
Ala Val Ala Asn Val Asn Asp Val His Gln Ala Thr Phe Leu Leu Ala
            90                  95                  100
Gln Thr Thr Leu Arg Asn Val Leu Gly Thr Gln Thr Leu Ser Gln Ile
            105                 110                 115
Leu Ala Gly Arg Glu Glu Ile Ala His Ser Ile Gln Thr Leu Leu Asp
            120                 125                 130
Asp Ala Thr Glu Leu Trp Gly Ile Arg Val Ala Arg Val Glu Ile Lys
135                 140                 145                 150
Asp Val Arg Ile Pro Val Gln Leu Gln Arg Ser Met Ala Ala Glu Ala
            155                 160                 165
Glu Ala Thr Arg Glu Ala Arg Ala Lys Val Leu Ala Ala Glu Gly Glu
            170                 175                 180
Met Ser Ala Ser Lys Ser Leu Lys Ser Ala Ser Met Val Leu Ala Glu
            185                 190                 195
Ser Pro Ile Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu Ser Thr Val
            200                 205                 210
Ala Thr Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro Met Asn Ile
215                 220                 225                 230
Leu Glu Gly Ile Gly Gly Val Ser Tyr Asp Asn His Lys Lys Leu Pro
            235                 240                 245
Asn Lys Ala
```

```
<210> 108
<211> 67
```

137 266

<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -26..-1
<400> 108
```
Met Ser Thr Trp Leu Leu Leu Ile Ala Leu Lys Thr Leu Ile Thr Trp
    -25                 -20                 -15
Val Ser Leu Phe Ile Asp Cys Val Met Thr Arg Lys Leu Thr Asn Cys
-10                  -5              1               5
Asn Ala Arg Glu Thr Ile Lys Gly Ile Gln Lys Arg Glu Ala Ser Asn
            10                  15                  20
Cys Phe Ala Ile Arg His Phe Glu Asn Lys Phe Ala Val Glu Thr Leu
            25                  30                  35
Ile Cys Ser
    40
```

<210> 109
<211> 127
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -63..-1
<400> 109
```
Met Ser Ala Ala Gly Ala Arg Gly Leu Arg Ala Thr Tyr His Arg Leu
            -60                 -55                 -50
Leu Asp Lys Val Glu Leu Met Leu Pro Glu Lys Leu Arg Pro Leu Tyr
            -45                 -40                 -35
Asn His Pro Ala Gly Pro Arg Thr Val Phe Phe Trp Ala Pro Ile Met
        -30                 -25                 -20
Lys Trp Gly Leu Val Cys Ala Gly Leu Ala Asp Met Ala Arg Pro Ala
-15                 -10                 -5                  1
Glu Lys Leu Ser Thr Ala Gln Ser Ala Val Leu Met Ala Thr Gly Phe
            5                   10                  15
Ile Trp Ser Arg Tyr Ser Leu Val Ile Ile Pro Lys Asn Trp Ser Leu
            20                  25                  30
Phe Ala Val Asn Phe Phe Val Gly Ala Ala Gly Ala Ser Gln Leu Phe
    35                  40                  45
Arg Ile Trp Arg Tyr Asn Gln Glu Leu Lys Ala Lys Ala His Lys
50                  55                  60
```

<210> 110
<211> 97
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -20..-1
<220>
<221> UNSURE
<222> 53
<223> Xaa = any one of the twenty amino acids
<400> 110
```
Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly Ala Leu Leu Gly
-20                 -15                 -10                 -5
Thr Ala Trp Ala Arg Arg Ser Arg Asp Leu His Cys Gly Ala Cys Arg
                1               5                   10
```

267

Ala Leu Val Asp Glu Leu Glu Trp Glu Ile Ala Gln Val Asp Pro Lys
        15   -   20   25
Lys Thr Ile Gln Met Gly Ser Phe Arg Ile Asn Pro Asp Gly Ser Gln
    30   35   40
Ser Val Val Glu Val Thr Val Thr Xaa Ser Pro Lys Thr Lys Val Ala
45   50   55   60
His Ser Gly Phe Trp Met Lys Ile Arg Leu Leu Lys Lys Gly Pro Trp
                65   70   75
Ser

<210> 111
<211> 86
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -20..-1
<400> 111
Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly Ala Leu Leu Gly
-20   -15   -10   -5
Thr Ala Trp Ala Arg Arg Ser Gln Asp Leu His Cys Gly Ala Cys Arg
                1   5   10
Ala Leu Val Asp Glu Thr Arg Met Gly Asn Cys Pro Gly Gly Pro Gln
        15   20   25
Glu Asp His Ser Asp Gly Ile Phe Pro Asp Gln Ser Arg Trp Gln Pro
    30   35   40
Val Ser Gly Gly Gly Ala Leu Cys Pro Leu Arg Gly Pro Pro His Arg
45   50   55   60
Ala Ala Gly Gly Asp Met
                65

<210> 112
<211> 71
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -25..-1
<400> 112
Met Pro Ala Gly Val Pro Met Ser Thr Tyr Leu Lys Met Phe Ala Ala
-25   -20   -15   -10
Ser Leu Leu Ala Met Cys Ala Gly Ala Glu Val Val His Arg Tyr Tyr
                -5   1   5
Arg Pro Asp Leu Thr Ile Pro Glu Ile Pro Pro Lys Arg Gly Glu Leu
        10   15   20
Lys Thr Glu Leu Leu Gly Leu Lys Glu Arg Lys His Lys Pro Gln Val
    25   30   35
Ser Gln Gln Glu Glu Leu Lys
40   45

<210> 113
<211> 60
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 113

139

```
Met Asp Gly His Trp Ser Ala Ala Phe Ser Ala Leu Thr Val Thr Ala
        -40             -35             -30
Met Ser Ser Trp Ala Arg Arg Arg Ser Ser Ser Ser Arg Arg Ile Pro
    -25             -20             -15
Ser Leu Pro Gly Ser Pro Val Cys Trp Ala Trp Pro Trp Tyr Pro Asp
-10             -5              1               5
Thr Thr Ser Phe Pro Leu Arg Cys Arg Gly Arg Val
            10              15
```

<210> 114
<211> 118
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -83..-1
<220>
<221> UNSURE
<222> 28,32
<223> Xaa = any one of the twenty amino acids
<400> 114

```
Met Leu Pro Val Gln Ser Phe Thr Leu Val Ala Gln Ala Gly Val Gln
        -80             -75             -70
Trp Arg His Leu Ser Ser Leu Gln Leu Leu Pro Pro Glu Phe Lys Gly
        -65             -60             -55
Phe Ser Cys Leu Ser Leu Pro Ser Ser Trp Asp Tyr Arg Arg Pro Pro
    -50             -45             -40
Pro Cys Pro Ala Gly Phe Phe Val Phe Leu Val Glu Thr Gly Leu His
-35             -30             -25             -20
His Val Gly Gln Ala Gly Leu Glu Leu Leu Thr Ser Cys Ser Pro Pro
            -15             -10             -5
Ala Ser Ala Ser Gln Ser Ala Ala Ile Thr Gly Val Ser His Val Pro
        1               5               10
Gly Lys Lys Lys Leu Leu Lys Val Glu Lys Lys Asn Leu Arg Xaa Leu
    15              20              25
Leu Thr Xaa Ile Lys Thr
30              35
```

<210> 115
<211> 76
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<220>
<221> UNSURE
<222> 22,43
<223> Xaa = any one of the twenty amino acids
<400> 115

```
Met Glu Leu Ile Ser Pro Thr Val Ile Ile Ile Leu Gly Cys Leu Ala
        -20             -15             -10
Leu Phe Leu Leu Leu Gln Arg Lys Asn Leu Arg Arg Pro Pro Cys Ile
    -5              1               5               10
Lys Gly Trp Ile Pro Trp Ile Gly Val Gly Phe Xaa Phe Gly Lys Ala
            15              20              25
Pro Leu Glu Phe Ile Glu Lys Ala Arg Ile Lys Val Cys Gly Arg Gly
            30              35              40
```

```
Xaa Arg Gly Leu Gln Arg Arg Gln Cys Phe Leu Phe
        45                  50


<210> 116
<211> 95
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -52..-1
<400> 116
Met Ala Glu Thr Lys Asp Ala Ala Gln Met Leu Val Thr Phe Lys Asp
        -50                 -45                 -40
Val Ala Val Thr Phe Thr Arg Glu Glu Trp Arg Gln Leu Asp Leu Ala
    -35                 -30                 -25
Gln Arg Thr Leu Tyr Arg Glu Val Met Leu Glu Thr Cys Gly Leu Leu
-20                 -15                 -10                     -5
Val Ser Leu Gly Gln Ser Ile Trp Leu His Ile Thr Glu Asn Gln Ile
                1                   5                   10
Lys Leu Ala Ser Pro Gly Arg Lys Phe Thr Asn Ser Pro Asp Glu Lys
        15                  20                  25
Pro Glu Val Trp Leu Ala Pro Gly Leu Phe Gly Ala Ala Ala Gln
    30                  35                  40


<210> 117
<211> 82
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<400> 117
Met Glu Leu Ile Ser Pro Thr Val Ile Ile Ile Leu Gly Cys Leu Ala
        -20                 -15                 -10
Leu Phe Leu Leu Leu Gln Arg Lys Asn Leu Arg Arg Pro Pro Cys Ile
    -5                  1                   5                   10
Lys Gly Trp Ile Pro Trp Ile Gly Val Gly Phe Glu Phe Gly Lys Ala
                15                  20                  25
Pro Leu Glu Phe Ile Glu Lys Ala Arg Ile Lys Tyr Gly Pro Ile Phe
                30                  35                  40
Thr Val Phe Ala Met Gly Asn Arg Met Thr Phe Val Thr Glu Glu Gly
        45                  50                  55
Arg Asn
    60


<210> 118
<211> 89
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -16..-1
<400> 118
Met Ile Ile Ser Leu Phe Ile Tyr Ile Phe Leu Thr Cys Ser Asn Thr
        -15                 -10                 -5
Ser Pro Ser Tyr Gln Gly Thr Gln Leu Gly Leu Gly Leu Pro Ser Ala
1                   5                   10                  15
Gln Trp Trp Pro Leu Thr Gly Arg Arg Met Gln Cys Cys Arg Leu Phe
```

141

```
                 20                      25                      30
Cys Phe Leu Leu Gln Asn Cys Leu Phe Pro Phe Pro Leu His Leu Ile
         35                      40                      45
Gln His Asp Pro Cys Glu Leu Val Leu Thr Ile Ser Trp Asp Trp Ala
     50                      55                      60
Glu Ala Gly Ala Ser Leu Tyr Ser Pro
65                      70


<210> 119
<211> 30
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -19..-1
<400> 119
Met Thr Met Ala Glu Cys Pro Thr Leu Cys Val Ser Ser Ser Pro Ala
                 -15                     -10                     -5
Leu Trp Ala Ala Ser Glu Thr Thr Asp Asp Val Cys Arg Glu
         1                       5                       10


<210> 120
<211> 115
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -103..-1
<400> 120
Met Val Ile Arg Val Tyr Ile Ala Ser Ser Ser Gly Ser Thr Ala Ile
                 -100                    -95                     -90
Lys Lys Lys Gln Gln Asp Val Leu Gly Phe Leu Glu Ala Asn Lys Ile
         -85                     -80                     -75
Gly Phe Glu Glu Lys Asp Ile Ala Ala Asn Glu Glu Asn Arg Lys Trp
     -70                     -65                     -60
Met Arg Glu Asn Val Pro Glu Asn Ser Arg Pro Ala Thr Gly Asn Pro
-55                     -50                     -45                     -40
Leu Pro Pro Gln Ile Phe Asn Glu Ser Gln Tyr Arg Gly Asp Tyr Asp
                 -35                     -30                     -25
Ala Phe Phe Glu Ala Arg Glu Asn Asn Ala Val Tyr Ala Phe Leu Gly
             -20                     -15                     -10
Leu Thr Ala Pro Ser Gly Ser Lys Glu Ala Gly Arg Cys Lys Gln Ser
         -5                      1                       5
Ser Lys Pro
10


<210> 121
<211> 105
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -76..-1
<400> 121
Met Pro Leu Leu Cys Gln Ile Glu Met Glu Tyr Leu Leu Leu Lys Trp
         -75                     -70                     -65
Gln Met Thr Met Leu Gln Ser Met Leu Cys Asp Leu Val Ser Tyr Pro
-60                     -55                     -50                     -45
```

142

```
Leu Leu Pro Leu Gln Gln Thr Lys Glu Ala Asn Leu Asp Phe Pro Lys
            -40                 -35                 -30
Ile Lys Val Ser Ser Val Thr Ile Thr Pro Thr Arg Trp Phe Asn Leu
            -25                 -20                 -15
Ile Val Tyr Leu Trp Val Val Ser Phe Ile Ala Ser Ser Ser Ala Asn
            -10                 -5                  1
Thr Gly Leu Ile Val Ser Leu Glu Lys Glu Leu Ala Pro Leu Phe Glu
5                   10                  15                  20
Glu Leu Arg Gln Val Val Glu Val Ser
                25

<210> 122
<211> 93
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<400> 122
Met Lys Pro Val Leu Pro Leu Gln Phe Leu Val Val Phe Cys Leu Ala
            -20                 -15                 -10
Leu Gln Leu Val Pro Gly Ser Pro Lys Gln Arg Val Leu Lys Tyr Ile
    -5                  1                   5                   10
Leu Glu Pro Pro Pro Cys Ile Ser Ala Pro Glu Asn Cys Thr His Leu
                15                  20                  25
Cys Thr Met Gln Glu Asp Cys Glu Lys Gly Phe Gln Cys Cys Ser Ser
            30                  35                  40
Phe Cys Gly Ile Val Cys Ser Ser Glu Thr Phe Gln Lys Arg Asn Arg
            45                  50                  55
Ile Lys His Lys Gly Ser Glu Val Ile Met Pro Ala Asn
    60                  65                  70

<210> 123
<211> 109
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 123
Met His Ile Leu Gln Leu Leu Thr Thr Val Asp Asp Gly Ile Gln Ala
            -40                 -35                 -30
Ile Val His Cys Pro Asp Thr Gly Lys Asp Ile Trp Asn Leu Leu Phe
            -25                 -20                 -15
Asp Leu Val Cys His Glu Phe Cys Gln Ser Asp Asp Pro Ala Ile Ile
-10                 -5                  1                   5
Leu Gln Glu Gln Lys Thr Val Leu Ala Ser Val Phe Ser Val Leu Ser
            10                  15                  20
Ala Ile Tyr Ala Ser Gln Thr Glu Gln Glu Tyr Leu Lys Ile Glu Lys
            25                  30                  35
Val Asp Leu Pro Leu Ile Asp Ser Leu Ile Arg Val Leu Gln Asn Met
    40                  45                  50
Glu Gln Cys Gln Lys Lys Pro Glu Asn Ser Ala Gly Val
55                  60                  65

<210> 124
<211> 51
<212> PRT
```

143

<213> Homo sapiens
<220>
<221> SIGNAL
<222> -15..-1
<400> 124

```
Met Arg Leu Val Pro Leu Gly Gln Ser Phe Pro Leu Ser Glu Pro Arg
-15                 -10                 -5                      1
Cys Leu Gln Pro Val Lys Trp Asp His Asn His Cys Leu Thr Ser Leu
            5                   10                  15
Thr Val Val Val Arg Thr Glu Cys Val Glu Val Phe His Lys Leu Trp
            20                  25                  30
Met Leu Val
    35
```

<210> 125
<211> 56
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -27..-1
<400> 125

```
Met Asn Arg Val Pro Ala Asp Ser Pro Asn Met Cys Leu Ile Cys Leu
        -25                 -20                 -15
Leu Ser Tyr Ile Ala Leu Gly Ala Ile His Ala Lys Ile Cys Arg Arg
        -10                 -5                  1                   5
Ala Phe Gln Glu Glu Gly Arg Ala Asn Ala Lys Thr Gly Val Arg Ala
                10                  15                  20
Trp Cys Ile Gln Pro Trp Ala Lys
            25
```

<210> 126
<211> 162
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -21..-1
<400> 126

```
Met Leu Gln Thr Ser Asn Tyr Ser Leu Val Leu Ser Leu Gln Phe Leu
    -20                 -15                 -10
Leu Leu Ser Tyr Asp Leu Phe Val Asn Ser Phe Ser Glu Leu Leu Gln
-5                  1                   5                   10
Lys Thr Pro Val Ile Gln Leu Val Leu Phe Ile Ile Gln Asp Ile Ala
            15                  20                  25
Val Leu Phe Asn Ile Ile Ile Phe Leu Met Phe Phe Asn Thr Ser
            30                  35                  40
Val Phe Gln Ala Gly Leu Val Asn Leu Leu Phe His Lys Phe Lys Gly
    45                  50                  55
Thr Ile Ile Leu Thr Ala Val Tyr Phe Ala Leu Ser Ile Ser Leu His
60                  65                  70                  75
Val Trp Val Met Asn Leu Arg Trp Lys Asn Ser Asn Ser Phe Ile Trp
                80                  85                  90
Thr Asp Gly Leu Gln Met Leu Phe Val Phe Gln Arg Leu Ala Ala Val
                95                  100                 105
Leu Tyr Cys Tyr Phe Tyr Lys Arg Thr Ala Val Arg Leu Gly Asp Pro
            110                 115                 120
His Phe Tyr Gln Asp Ser Leu Trp Leu Arg Lys Glu Phe Met Gln Val
```

144

```
        125                    130                      135
Arg Arg
140


<210> 127
<211> 126
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -68..-1
<400> 127
Met Ala Ser Ala Ser Ala Arg Gly Asn Gln Asp Lys Asp Ala His Phe
        -65                    -60                      -55
Pro Pro Pro Ser Lys Gln Ser Leu Leu Phe Cys Pro Lys Ser Lys Leu
        -50                    -45                      -40
His Ile His Arg Ala Glu Ile Ser Lys Ile Met Arg Glu Cys Gln Glu
    -35                    -30                      -25
Glu Ser Phe Trp Lys Arg Ala Leu Pro Phe Ser Leu Val Ser Met Leu
-20                    -15                      -10                 -5
Val Thr Gln Gly Leu Val Tyr Gln Gly Tyr Leu Ala Ala Asn Ser Arg
                  1                 5                      10
Phe Gly Ser Leu Pro Lys Val Ala Leu Ala Gly Leu Leu Gly Phe Gly
        15                     20                      25
Leu Gly Lys Val Ser Tyr Ile Gly Val Cys Gln Ser Lys Phe His Phe
    30                     35                      40
Phe Glu Asp Gln Leu Arg Gly Ala Gly Phe Gly Pro Thr Ala
45                     50                      55


<210> 128
<211> 140
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -40..-1
<400> 128
Met Thr Ser Met Thr Gln Ser Leu Arg Glu Val Ile Lys Ala Met Thr
-40                    -35                      -30                 -25
Lys Ala Arg Asn Phe Glu Arg Val Leu Gly Lys Ile Thr Leu Val Ser
                  -20                    -15                      -10
Ala Ala Pro Gly Lys Val Ile Cys Glu Met Lys Val Glu Glu Glu His
              -5                     1                 5
Thr Asn Ala Ile Gly Thr Leu His Gly Gly Leu Thr Ala Thr Leu Val
        10                     15                      20
Asp Asn Ile Ser Thr Met Ala Leu Leu Cys Thr Glu Arg Gly Ala Pro
25                     30                      35                  40
Gly Val Ser Val Asp Met Asn Ile Thr Tyr Met Ser Pro Ala Lys Leu
              45                     50                      55
Gly Glu Asp Ile Val Ile Thr Ala His Val Leu Lys Gln Gly Lys Thr
              60                     65                      70
Leu Ala Phe Thr Ser Val Gly Leu Thr Asn Lys Ala Thr Gly Lys Leu
        75                     80                      85
Ile Ala Gln Gly Arg His Thr Lys His Leu Gly Asn
    90                     95                      100


<210> 129
<211> 43
```

<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -24..-1
<400> 129
Met Gln Cys Phe Ser Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu
              -20   -15   -10
Leu Ile Phe Leu Cys Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp Ser
              -5   1   5
Pro Tyr Phe Lys Met His Lys Pro Val Thr Met
    10   15

<210> 130
<211> 69
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -21..-1
<400> 130
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
              -20   -15   -10
Ile Trp Thr Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr
-5   1   5   10
Leu His His Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr
              15   20   25
Val Ala Pro Glu Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala
              30   35   40
Val Leu Cys Gln Lys
    45

<210> 131
<211> 78
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -19..-1
<400> 131
Met Ser Pro Gly Ser Ala Leu Ala Leu Leu Trp Ser Leu Pro Ala Ser
              -15   ,   -10   -5
Asp Leu Gly Arg Ser Val Ile Ala Gly Leu Trp Pro His Thr Gly Val
              1   5   10
Leu Ile His Leu Glu Thr Ser Gln Ser Phe Leu Gln Gly Gln Leu Thr
    15   20   25
Lys Ser Ile Phe Pro Leu Cys Cys Thr Ser Leu Phe Cys Val Cys Val
30   35   40   45
Val Thr Val Gly Gly Gly Arg Val Gly Ser Thr Phe Val Ala
              50   55

<210> 132
<211> 80
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -47..-1

<400> 132

```
Met Arg Leu Pro Pro Ala Leu Pro Ser Gly Tyr Thr Asp Ser Thr Ala
        -45                 -40                 -35
Leu Glu Gly Leu Val Tyr Tyr Leu Asn Gln Lys Leu Leu Phe Ser Ser
        -30                 -25                 -20
Pro Ala Ser Ala Leu Leu Phe Phe Ala Arg Pro Cys Val Phe Cys Phe
-15                 -10                 -5                  1
Lys Ala Ser Lys Met Gly Pro Gln Phe Glu Asn Tyr Pro Thr Phe Pro
            5                   10                  15
Thr Tyr Ser Pro Leu Pro Ile Ile Pro Phe Gln Leu His Gly Arg Phe
        20                  25                  30
```

<210> 133
<211> 53
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 133

```
Met Asp Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys Glu Asn Phe
        -40                 -35                 -30
Val Gly Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp Ile Leu Phe
        -25                 -20                 -15
Ser Leu Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile Ser Ile Trp
-10                 -5                  1                   5
Met Cys Leu Lys Ile
            10
```

<210> 134
<211> 1053
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 131..169
<223> Von Heijne matrix
      score 4.19999980926514
      seq MLAVSLTVPLLGA/MM
<220>
<221> polyA_site
<222> 1042..1053
<400> 134

```
gagcgagtcg gacgggctgc gacagcgccg gcccctgcgg ccgcaggtcg tcacagacga      60
tgatggccag gccccggagg ctaaggacgg cagctccttt agcggcagag ttttccgagt     120
gaccttcttg atg ctg gct gtt tct ctc acc gtt ccc ctg ctt gga gcc        169
            Met Leu Ala Val Ser Leu Thr Val Pro Leu Leu Gly Ala
                    -10                 -5
atg atg ctg ctg gaa tct cct ata gat cca cag cct ctc agc ttc aaa       217
Met Met Leu Leu Glu Ser Pro Ile Asp Pro Gln Pro Leu Ser Phe Lys
1               5                   10                  15
gaa ccc ccg ctc ttg ctt ggt gtt ctg cat cca aat acg aag ctg cga       265
Glu Pro Pro Leu Leu Leu Gly Val Leu His Pro Asn Thr Lys Leu Arg
            20                  25                  30
cag gca gaa agg ctg ttt gaa aat caa ctt gtt gga ccg gag tcc ata       313
Gln Ala Glu Arg Leu Phe Glu Asn Gln Leu Val Gly Pro Glu Ser Ile
        35                  40                  45
gca cat att ggg gat gtg atg ttt act ggg aca gca gat ggc cgg gtc       361
```

```
Ala His Ile Gly Asp Val Met Phe Thr Gly Thr Ala Asp Gly Arg Val
    50                  55                  60
gta aaa ctt gaa aat ggt gaa ata gag acc att gcc cgg ttt ggt tcg     409
Val Lys Leu Glu Asn Gly Glu Ile Glu Thr Ile Ala Arg Phe Gly Ser
65                  70                  75                  80
ggc cct tgc aaa acc cga ggt gat gag cct gtg tgt ggg aga ccc ctg     457
Gly Pro Cys Lys Thr Arg Gly Asp Glu Pro Val Cys Gly Arg Pro Leu
                85                  90                  95
ggt atc cgt gca ggg ccc aat ggg act ctc ttt gtg gcc gat gca tac     505
Gly Ile Arg Ala Gly Pro Asn Gly Thr Leu Phe Val Ala Asp Ala Tyr
                100                 105                 110
aag gga cta ttt gaa gta aat ccc tgg aaa cgt gaa gtg aaa ctg ctg     553
Lys Gly Leu Phe Glu Val Asn Pro Trp Lys Arg Glu Val Lys Leu Leu
                115                 120                 125
ctg tcc tcc gag aca ccc att gag ggg aag aac atg tcc ttt gtg aat     601
Leu Ser Ser Glu Thr Pro Ile Glu Gly Lys Asn Met Ser Phe Val Asn
    130                 135                 140
gat ctt aca gtc act cag gat ggg agg aag att tat ttc acc gat tct     649
Asp Leu Thr Val Thr Gln Asp Gly Arg Lys Ile Tyr Phe Thr Asp Ser
145                 150                 155                 160
agc agc aaa tgg caa aga cga gac tac ctg ctt ctg gtg atg gag ggc     697
Ser Ser Lys Trp Gln Arg Arg Asp Tyr Leu Leu Leu Val Met Glu Gly
                165                 170                 175
aca gat gac ggg cgc ctg ctg gag tat gat act gtg acc agg gaa gta     745
Thr Asp Asp Gly Arg Leu Leu Glu Tyr Asp Thr Val Thr Arg Glu Val
                180                 185                 190
aaa gtt tta ttg gac cag ctg cgg ttc ccg aat gga gtc cag ctg tct     793
Lys Val Leu Leu Asp Gln Leu Arg Phe Pro Asn Gly Val Gln Leu Ser
                195                 200                 205
cct gca gaa gac ttt gtc ctg gtg gca gaa aca acc atg gcc agg ata     841
Pro Ala Glu Asp Phe Val Leu Val Ala Glu Thr Thr Met Ala Arg Ile
    210                 215                 220
cga aga gtc tac gtt tct ggc ctg atg aag ggc ggg gct gat ctg ttt     889
Arg Arg Val Tyr Val Ser Gly Leu Met Lys Gly Gly Ala Asp Leu Phe
225                 230                 235                 240
gtg gag aac atg cct gga ttt cca gac aac atc cgg ccc agc agc tct     937
Val Glu Asn Met Pro Gly Phe Pro Asp Asn Ile Arg Pro Ser Ser Ser
                245                 250                 255
ggg ggg tac tgg gtg ggc atg tcg acc atc cgc cct aac cct ggg ttt     985
Gly Gly Tyr Trp Val Gly Met Ser Thr Ile Arg Pro Asn Pro Gly Phe
                260                 265                 270
tcc atg ctg gat ttc tta tct gag aga ccc tgg att aaa agg atg att    1033
Ser Met Leu Asp Phe Leu Ser Glu Arg Pro Trp Ile Lys Arg Met Ile
                275                 280                 285
ttt aag gta aaaaaaaaaa a                                           1053
Phe Lys Val
    290

<210> 135
<211> 675
<212> DNA
<213> Homo sapiens
<220>
<221> polyA_signal
<222> 638..643
<220>
<221> polyA_site
<222> 662..675
```

148

<400> 135

```
accgaacagg aacagcacaa cctgggaccc agacatgcag tacctctacg caaagtaaaa        60
gtagcagtgg ttcagcacac tttggtatgt tgactgtta atg atg tac gtt tct       114
                                              Met Met Tyr Val Ser
                                                1               5
ata gaa atg tca ggt cca acc att tcc cat ttg ttc gac tat gtg gtc       162
Ile Glu Met Ser Gly Pro Thr Ile Ser His Leu Phe Asp Tyr Val Val
            10               15                  20
tgt tac att tat ggc tta aag tcc ttt tct ctt aaa cag tta aaa aaa       210
Cys Tyr Ile Tyr Gly Leu Lys Ser Phe Ser Leu Lys Gln Leu Lys Lys
            25               30                  35
aaa tct tgg tct aag tat tta ttt gaa tcc tgt tgc tat agg agt ttg       258
Lys Ser Trp Ser Lys Tyr Leu Phe Glu Ser Cys Cys Tyr Arg Ser Leu
            40               45                  50
tat gtg tgt gtc ttc att taaacatacc tgcatacaaa gatggtttat             306
Tyr Val Cys Val Phe Ile
            55
ttctatttaa tatgtgacat ttgtttcctg gatatagtcc gtgaaccaca agatttatca       366
tattttttcaa taatatgaga agaaaatggg ccgtaaattg ttaaccattt tatgttcaga       426
tatttctcta gttttttacct agtttgcttt aacatagaga ccagcaagtg aatatatatg       486
cataacctta tatgttgaca caataattca gaataatttg ttaaagataa actaattttt       546
cagagaagaa catttaaagg gttaatattt ttgaaacgtt ttcagataat atctatttga       606
ttattgtggc ttctatttga aatgtgtcta aaataaatgc tgtttatttta aaatgaaaaa       666
aaaaaaaaa                                                             675
```

<210> 136
<211> 1112
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 111..194
<223> Von Heijne matrix
      score 4.80000019073486
      seq GVLLEPFVHQVGG/HS
<220>
<221> polyA_signal
<222> 1080..1085
<220>
<221> polyA_site
<222> 1101..1112
<400> 136

```
ccgagagaga ctacacggta ctgggacaca cggacaaaca acagacagaa gacgtactgg        60
ccgctggact ccgctgcctc ccccatctcc ccgccatctg cgcccggagg atg agc       116
                                                          Met Ser
cca gcc ttc agg gcc atg gat gtg gag ccc cgc gcc aaa ggc gtc ctt       164
Pro Ala Phe Arg Ala Met Asp Val Glu Pro Arg Ala Lys Gly Val Leu
        -25               -20               -15
ctg gag ccc ttt gtc cac cag gtc ggg ggg cac tca tgc gtg ctc cgc       212
Leu Glu Pro Phe Val His Gln Val Gly Gly His Ser Cys Val Leu Arg
-10               -5                  1                  5
ttc aat gag aca acc ctg tgc aag ccc ctg gtc cca agg gaa cat cag       260
Phe Asn Glu Thr Thr Leu Cys Lys Pro Leu Val Pro Arg Glu His Gln
            10               15                  20
ttc tac gag acc ctc cct gct gag atg cgc aaa ttc tct ccc cag tac       308
Phe Tyr Glu Thr Leu Pro Ala Glu Met Arg Lys Phe Ser Pro Gln Tyr
            25               30                  35
aaa gga caa agc caa agg ccc ctt gtt agc tgg cca tcc ctg ccc cat       356
```

149

```
Lys Gly Gln Ser Gln Arg Pro Leu Val Ser Trp Pro Ser Leu Pro His
    40                  45                  50
ttt ttc ccc tgg tcc ttt ccc ctg tgg cca cag gga agt gtg gcc          401
Phe Phe Pro Trp Ser Phe Pro Leu Trp Pro Gln Gly Ser Val Ala
55                  60                  65
tgaataccc accccggctc ctctgcaccc agagctgggg gccacctcag aagtgtcatc     461
tctctctgag cacgcattcc cctgcagcag tcgaggactg agcagattga gtgatgctgg    521
ggcagagagg cctgagagga aaggtgttca gccagtcgtt tgtaaggcgc tcgtcggcac    581
ctgctgaaac gccccacct gacagcccca tcctcaaaga ctgtcttaat tactcatggc     641
aggttctaga gacttaaggg gaaaagctgc tttcaaggcc accacatgtc tgtgctcccc    701
aaccagctct atctgccttg tgttcatttt gttattttgt gacgtgagac agcaaagacc    761
aataaaaaca tattttataa gaacaaaagg cctgggtgcc tacccgtgtg ggggcactgt    821
gggaagcctt tgctagggtg tcttgtgctg tgtggtttgt tttgtttgcc cctttatttt    881
gctttgctta cccagtcttc ccttactctt ggatgcttct taaccctcag gcaaacctgt    941
gttccccctg tattcaggct ctgctttaaa gcaagccatg aggctgttgg agtttctgtt    1001
tagggcatta aaaattcccg caaactataa agagcaatgt tttcagtctt ttaggattag    1061
aagaattaca taaaaattaa taaacatttt caatgatgga aaaaaaaaaa a            1112

<210> 137
<211> 547
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 359..454
<223> Von Heijne matrix
      score 4
      seq FSFMLLGMGGCLP/GF
<220>
<221> polyA_site
<222> 536..547
<400> 137
ctggggagcc ctgcctaaga ctcatgctac aagaagttaa ataagtttcc cgaagtcaca     60
cagctagcct ctcatccctt ttctactgag aggaagtgga atgcactccg acaaggataa    120
ggttttattg tgagctggcc ttggaattaa accaccacca acacactttt ggattatcag    180
aaggtggaag gagtgcaaaa atgtcattcc catgcttgtc tgccaggcaa cctggtgtcc    240
attctttatg acgcctttcc tgaatcacag gtgcattggg gtgcttcctc ctccccagga    300
ctcccaccca actttgtgaa cacaacccac ttagaggagt tatctcagca cattatga     358
atg ttg ggg acc acg ggc ctc ggg aca cag ggt cct tcc cag cag gct      406
Met Leu Gly Thr Thr Gly Leu Gly Thr Gln Gly Pro Ser Gln Gln Ala
        -30                 -25                 -20
ctg ggc ttt ttc tcc ttt atg tta ctt gga atg ggc ggg tgc ctg cct      454
Leu Gly Phe Phe Ser Phe Met Leu Leu Gly Met Gly Gly Cys Leu Pro
    -15                 -10                 -5
gga ttc ctg cta cag cct ccc aat cga tct cct act ttg cct gca tcc      502
Gly Phe Leu Leu Gln Pro Pro Asn Arg Ser Pro Thr Leu Pro Ala Ser
1               5                   10                  15
acc ttt gcc cat taaagtcaat ctccaccca taaaaaaaaa aaa                  547
Thr Phe Ala His
            20

<210> 138
<211> 1198
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 26..316
```

<223> Von Heijne matrix
      score 4
      seq RLPLVVSFIASSS/AN
<220>
<221> polyA_signal
<222> 1164..1169
<220>
<221> polyA_site
<222> 1187..1198
<400> 138
```
atcctgcgaa agaaggggt tcatc atg gcg gat gac cta aag cga ttc ttg          52
                              Met Ala Asp Asp Leu Lys Arg Phe Leu
                                  -95                 -90
tat aaa aag tta cca agt gtt gaa ggg ctc cat gcc att gtt gtg tca         100
Tyr Lys Lys Leu Pro Ser Val Glu Gly Leu His Ala Ile Val Val Ser
        -85             -80                 -75
gat aga gat gga gta cct gtt gtt aaa gtg gca aat gac aat gct cca        148
Asp Arg Asp Gly Val Pro Val Val Lys Val Ala Asn Asp Asn Ala Pro
        -70             -65                 -60
gag cat gct ttg cga cct ggt ttc tta tcc act ttt gcc ctt gca aca        196
Glu His Ala Leu Arg Pro Gly Phe Leu Ser Thr Phe Ala Leu Ala Thr
    -55             -50                 -45
gac caa gga agc aaa ctt gga ctt tcc aaa aat aaa agt atc atc tgt        244
Asp Gln Gly Ser Lys Leu Gly Leu Ser Lys Asn Lys Ser Ile Ile Cys
-40             -35                 -30                 -25
tac tat aac acc tac cag gtg gtt caa ttt aat cgt tta cct ttg gtg        292
Tyr Tyr Asn Thr Tyr Gln Val Val Gln Phe Asn Arg Leu Pro Leu Val
            -20             -15                 -10
gtg agt ttc ata gcc agc agc agt gcc aat aca gga cta att gtc agc        340
Val Ser Phe Ile Ala Ser Ser Ser Ala Asn Thr Gly Leu Ile Val Ser
            -5              1                   5
cta gaa aag gag ctt gct cca ttg ttt gaa gaa ctg aga caa gtt gtg        388
Leu Glu Lys Glu Leu Ala Pro Leu Phe Glu Glu Leu Arg Gln Val Val
    10              15                  20
gaa gtt tct taatctgaca gtggtttcag tgtgtacctt atcttcatta              437
Glu Val Ser
25
taacaacaca atatcaatcc agcaatcttt agactacaat aatgctttta tccatgtgct     497
caagaaaggg cccctttttc caacttatac taaagagcta gcatatagat gtaatttata     557
gatagatcag ttgctatatt ttctggtgta gggtctttct tatttagtga gatctaggga     617
taccacagaa atggttcagt ctatcacagc tcccatggag ttagtctggt caccagatat     677
ggatgagaga ttctattcag tggattagaa tcaaactggt acattgatcc acttgagccg     737
ttaagtgctg ccaattgtac aatatgccca ggcttgcaga ataaagccaa cttttttattg    797
tgaataataa taaggacata tttttcttca gattatgttt tatttctttg cattgagtga     857
ggtacataaa atggcttggt aaaagtaata aaatcagtac aatcactaac tttccttgt      917
acatattatt ttgcagtata gatgaatatt actaatcagt ttgattattc tcagagggtg     977
ctgctcttta atgaaaatga aaattatagc taatgttttt tcctcaaact ctgctttctg    1037
taaccaatca gtgtttttaat gtttgtgtgt tcttcataaa atttaaatac aattcgttat   1097
tctgtttcca atgttagtat gtatgtaaac atgatagtac agccattttt ttcatatgtg    1157
agtaaaaata aaatagtatt tttaaaagta aaaaaaaaa a                         1198
```
<210> 139
<211> 1400
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 36..107

151

<223> Von Heijne matrix
      score 5.69999980926514
      seq ILGLLGLLGTLVA/ML
<220>
<221> polyA_signal
<222> 1302..1307
<220>
<221> polyA_site
<222> 1389..1400
<400> 139

```
cagtccctga agacgcttct actgagaggt ctgcc atg gcc tct ctt ggc ctc          53
                                        Met Ala Ser Leu Gly Leu
                                                            -20

caa ctt gtg ggc tac atc cta ggc ctt ctg ggg ctt ttg ggc aca ctg        101
Gln Leu Val Gly Tyr Ile Leu Gly Leu Leu Gly Leu Leu Gly Thr Leu
        -15                 -10                 -5

gtt gcc atg ctg ctc ccc agc tgg aaa aca agt tct tat gtc ggt gcc        149
Val Ala Met Leu Leu Pro Ser Trp Lys Thr Ser Ser Tyr Val Gly Ala
         1               5                   10

agc att gtg aca gca gtt ggc ttc tcc aag ggc ctc tgg atg gaa tgt        197
Ser Ile Val Thr Ala Val Gly Phe Ser Lys Gly Leu Trp Met Glu Cys
15                  20                  25                  30

gcc aca cac agc aca ggc atc acc cag tgt gac atc tat agc acc ctt        245
Ala Thr His Ser Thr Gly Ile Thr Gln Cys Asp Ile Tyr Ser Thr Leu
            35                  40                  45

ctg ggc ctg ccc gct gac atc cag gct gcc cag gcc atg atg gtg aca        293
Leu Gly Leu Pro Ala Asp Ile Gln Ala Ala Gln Ala Met Met Val Thr
            50                  55                  60

tcc agt gca atc tcc tcc ctg gcc tgc att atc tct gtg gtg ggc atg        341
Ser Ser Ala Ile Ser Ser Leu Ala Cys Ile Ile Ser Val Val Gly Met
            65                  70                  75

aga tgc aca gtc ttc tgc cag gaa tcc cga gcc aaa gac aga gtg gcg        389
Arg Cys Thr Val Phe Cys Gln Glu Ser Arg Ala Lys Asp Arg Val Ala
        80                  85                  90

gta gca ggt gga gtc ttt ttc atc ctt gga ggc ctc ctg gga ttc att        437
Val Ala Gly Gly Val Phe Phe Ile Leu Gly Gly Leu Leu Gly Phe Ile
95                  100                 105                 110

cct gtt gcc tgg aat ctt cat ggg atc cta cgg gac ttc tac tca cca        485
Pro Val Ala Trp Asn Leu His Gly Ile Leu Arg Asp Phe Tyr Ser Pro
                115                 120                 125

ctg gtg cct gac agc atg aaa ttt gag att gga gag gct ctt tac ttg        533
Leu Val Pro Asp Ser Met Lys Phe Glu Ile Gly Glu Ala Leu Tyr Leu
            130                 135                 140

ggc att att tct tcc ctg ttc tcc ctg ata gct gga atc atc ctc tgc        581
Gly Ile Ile Ser Ser Leu Phe Ser Leu Ile Ala Gly Ile Ile Leu Cys
            145                 150                 155

ttt tcc tgc tca tcc cag aga aat cgc tcc aac tac tac gat gcc tac        629
Phe Ser Cys Ser Ser Gln Arg Asn Arg Ser Asn Tyr Tyr Asp Ala Tyr
            160                 165                 170

caa gcc caa cct ctt gcc aca agg agc tct cca agg cct ggt caa cct        677
Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser Pro Arg Pro Gly Gln Pro
175                 180                 185                 190

ccc aaa gtc aag agt gag ttc aat tcc tac agc ctg aca ggg tat gtg        725
Pro Lys Val Lys Ser Glu Phe Asn Ser Tyr Ser Leu Thr Gly Tyr Val
                195                 200                 205

tgaagaacca ggggccagag ctggggggtg gctgggtctg tgaaaaacag tggacagcac      785
cccgagggcc acaggtgagg gacactacca ctggatcgtg tcagaaggtg ctgctgaggg      845
tagactgact ttggccattg gattgagcaa aggcagaaat ggggctagt gtaacagcat       905
```

152

```
gcaggttgaa ttgccaagga tgctcgccat gccagccttt ctgttttcct caccttgctg    965
ctcccctgcc ctaagtcccc aaccctcaac ttgaaacccc attcccttaa gccaggactc   1025
agaggatccc tttgccctct ggtttacctg ggactccatc cccaaaccca ctaatcacat   1085
cccactgact gaccctctgt gatcaaagac cctccctctg ctgaggttg gctcttagct     1145
cattgctggg gatgggaagg agaagcagtg gcttttgtgg gcattgctct aacctacttc   1205
tcaagcttcc ctccaaagaa actgattggc cctggaacct ccatcccact cttgttatga   1265
ctccacagtg tccagactaa tttgtgcatg aactgaaata aaaccatcct acggtatcca   1325
gggaacagaa agcaggatgc aggatgggag gacaggaagg cagcctggga catttaaaaa   1385
aataaaaaaa aaaaa                                                      1400
```

<210> 140
<211> 538
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 35..130
<223> Von Heijne matrix
     score 8
     seq VPMLLLIAGGSFG/LR
<220>
<221> polyA_signal
<222> 505..510
<220>
<221> polyA_site
<222> 526..538
<400> 140

```
gcttggagtt ctgagccgat ggaggagttc actc atg ttt gca ctc gcg gtg atg    55
                                      Met Phe Ala Leu Ala Val Met
                                                              -30

cgt gct ttt cgc aag aac aag act ctc ggc tat gga gtc ccc atg ttg     103
Arg Ala Phe Arg Lys Asn Lys Thr Leu Gly Tyr Gly Val Pro Met Leu
-25               -20               -15                      -10

ttg ctg att gct gga ggt tct ttt ggt ctt cgt gag ttt tct caa atc     151
Leu Leu Ile Ala Gly Gly Ser Phe Gly Leu Arg Glu Phe Ser Gln Ile
                  -5               1               5

cga tat gat gct gtg aag agt aaa atg gat cct gag ctt gaa aaa aaa     199
Arg Tyr Asp Ala Val Lys Ser Lys Met Asp Pro Glu Leu Glu Lys Lys
          10              15              20

ccg aaa gag aat aaa ata tct tta gag tcg gaa tat gag gga agt atc     247
Pro Lys Glu Asn Lys Ile Ser Leu Glu Ser Glu Tyr Glu Gly Ser Ile
    25              30              35

tgt tgaagggcta ctatctttcc ttggcccttc tcccttgttg ggactcaatc          300
Cys
40

tccagactat ctccccagag aatcttgtca aggcttggct ttaagctttg ttgggaaaat   360
caaagactcc aagtttgatg actggaagaa tattcgagga cccaggcctt gggaagatcc   420
tgacctcctc caaggaagaa atccagaaag ccttaagact aagacaactt gactctgctg   480
attctttttt cctttttttt tttaaataaa aatactatta actggaaaaa aaaaaaaa     538
```

<210> 141
<211> 1167
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 169..267
<223> Von Heijne matrix

153

```
<220>
<221> polyA_signal
<222> 1132..1137
<220>
<221> polyA_site
<222> 1155..1167
<400> 141
```
```
gtaggaacta ctgtcccaga gctgaggcaa ggggatttct caggtcattt ggagaacaag      60
tgctttagta gtagtttaaa gtagtaactg ctactgtatt tagtggggtg gaattcagaa     120
gaaatttgaa gaccagatca tgggtggtct gcatgtgaat gaacagga atg agc cag     177
                                                        Met Ser Gln
```
```
aca gcc tgg ctg tca ttg ctt tct tcc tcc cca ttt gga ccc ttc tct     225
Thr Ala Trp Leu Ser Leu Leu Ser Ser Ser Pro Phe Gly Pro Phe Ser
-30                 -25                 -20                 -15
gcc ctt aca ttt ttg ttt ctc cat cta cca cca tcc acc agt cta ttt     273
Ala Leu Thr Phe Leu Phe Leu His Leu Pro Pro Ser Thr Ser Leu Phe
              -10                 -5                  1
att aac tta gca aga gga caa ata aag ggc cct ctt ggc ttg att ttg     321
Ile Asn Leu Ala Arg Gly Gln Ile Lys Gly Pro Leu Gly Leu Ile Leu
          5                   10                  15
ctt ctt tct ttc tgt gga gga tat act aag tgc gac ttt gcc cta tcc     369
Leu Leu Ser Phe Cys Gly Gly Tyr Thr Lys Cys Asp Phe Ala Leu Ser
      20                  25                  30
tat ttg gaa atc cct aac aga att gag ttt tct att atg gat cca aaa     417
Tyr Leu Glu Ile Pro Asn Arg Ile Glu Phe Ser Ile Met Asp Pro Lys
35                  40                  45                  50
aga aaa aca aaa tgc taatgaagcc atcagtcaag ggtcacatgc caataaacaa     472
Arg Lys Thr Lys Cys
                  55
```
```
taaattttcc agaagaaatg aaatccaact agacaaataa agtagagctt atgaaatggt     532
tcagtaagga tgagcttgtt gttttttgtt ttgttttgtt ttgttttttt aaagacggag     592
tctcgctctg tcactcaggc tggagtgcag tggtatgatc ttggctcact gtaacctccg     652
cctcccgggt tcaagccatt ctcctgcctc agtctcctga gtagctggga ttgcaggtgc     712
gtgccaccat gcctggctaa ttttgtgtt tttggtagag acagggtttc accacgttgg     772
tcgggctggt ctcgggctcc tgacctcttg atccgcctgc cttggcctcc caaagtgatg     832
ggattacaga gtgtgagcca cgtgcctagc caaggatgag attttaaag tatgttccag     892
ttctgtgtca tggttggaag acagagtagg aaggatatgg aaaaggtcat ggggaagcag     952
aggtgattca tggctctgtg aatttgaggt gaatggttcc ttattgtcta ggccacttgt    1012
gaagaatatg agtcagttat gccagcctt ggaatttact tctctagctt acaatggacc    1072
ttttgaactg ggaaacacct tgtctgcatt cactttaaaa tgtcaaaact aattttttata    1132
ataaatgttt attttcacat cgaaaaaaaa aaaaa                               1167
```
```
<210> 142
<211> 730
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 143..238
<223> Von Heijne matrix
```
```
<220>
<221> polyA_signal
<222> 697..702
<220>
```

<221> polyA_site
<222> 721..730
<220>
<221> misc_feature
<222> 1,14,28,52
<223> n=a, g, c or t
<400> 142

```
nctttgcctt tctntccaca ggtgtccnct cccaggtcca actgcagact tngaattcgt    60
cttggtgaga gcgtgagctg ctgagatttg ggagtctgcg ctaggcccgc ttggagttct   120
gagccgatgg aagagttcac tc atg ttt gca ccc gcg gtg acg cgt gct ttt   172
                          Met Phe Ala Pro Ala Val Thr Arg Ala Phe
                                    -30                 -25
cgc aag aac aag act ctc ggc tat gga gtc ccc atg ttg ttg ctg att   220
Arg Lys Asn Lys Thr Leu Gly Tyr Gly Val Pro Met Leu Leu Leu Ile
          -20                 -15                 -10
gtt gga ggt tct ttt ggt ctt cgt gag ttt tct caa atc cga tat gat   268
Val Gly Gly Ser Phe Gly Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp
      -5                  1                   5                  10
gct gtg aag agt aaa atg gat cct gag ctt gaa aaa aaa ctg aaa gag   316
Ala Val Lys Ser Lys Met Asp Pro Glu Leu Glu Lys Lys Leu Lys Glu
              15                  20                  25
aat aaa ata tct tta gag tcg gaa tat gag aaa atc aaa gac tcc aag   364
Asn Lys Ile Ser Leu Glu Ser Glu Tyr Glu Lys Ile Lys Asp Ser Lys
              30                  35                  40
ttt gat gac tgg aag aat att cga gga ccc agg cct tgg gaa gat cct   412
Phe Asp Asp Trp Lys Asn Ile Arg Gly Pro Arg Pro Trp Glu Asp Pro
              45                  50                  55
gac ctc ctc caa gga aga aat cca gaa agc ctt aag act aag aca act   460
Asp Leu Leu Gln Gly Arg Asn Pro Glu Ser Leu Lys Thr Lys Thr Thr
      60                  65                  70
tgactctgct gattctcttt tccttttttt ttttaaataa aaatactatt aactggactt   520
cctaatatat acttctatca agtggaaagg aaattccagg cccatggaaa cttggatatg   580
ggtaatttga tgacaaataa tcttcactaa aggtcatgta caggtttta tacttcccag    640
ctattccatc tgtggatgaa agtaacaatg ttggccacgt atatttaca cctcgaaata    700
aaaaatgtga atactgctcc aaaaaaaaaa                                    730
```

<210> 143
<211> 1174
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 108..170
<223> Von Heijne matrix
      score 5.5
      seq SFLPSALVIWTSA/AF
<220>
<221> polyA_signal
<222> 1141..1146
<220>
<221> polyA_site
<222> 1161..1174
<400> 143

```
cacgttcctg ttgagtacac gttcctgttg atttacaaaa ggtgcaggta tgagcaggtc    60
tgaagactaa catttgtga agttgtaaaa cagaaaacct gttagaa atg tgg tgg   116
                                                    Met Trp Trp
                                                    -20
ttt cag caa ggc ctc agt ttc ctt cct tca gcc ctt gta att tgg aca   164
```

```
Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr
        -15             -10                 -5
tct gct gct ttc ata ttt tca tac att act gca gta aca ctc cac cat    212
Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His
        1                5                   10
ata gac ccg gct tta cct tat atc agt gac act ggt aca gta gct cca    260
Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro
15                  20                  25                  30
gaa aaa tgc tta ttt ggg gca atg cta aat att gcg gca gtt tta tgc    308
Glu Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys
            35                  40                  45
att gct acc att tat gtt cgt tat aag caa gtt cat gct ctg agt cct    356
Ile Ala Thr Ile Tyr Val Arg Tyr Lys Gln Val His Ala Leu Ser Pro
            50                  55                  60
gaa gag aac gtt atc atc aaa tta aac aag gct ggc ctt gta ctt gga    404
Glu Glu Asn Val Ile Ile Lys Leu Asn Lys Ala Gly Leu Val Leu Gly
        65                  70                  75
ata ctg agt tgt tta gga ctt tct att gtg gca aac ttc cag aaa aca    452
Ile Leu Ser Cys Leu Gly Leu Ser Ile Val Ala Asn Phe Gln Lys Thr
        80                  85                  90
acc ctt ttt gct gca cat gta agt gga gct gtg ctt acc ttt ggt atg    500
Thr Leu Phe Ala Ala His Val Ser Gly Ala Val Leu Thr Phe Gly Met
95                  100                 105                 110
ggc tca tta tat atg ttt gtt cag acc atc ctt tcc tac caa atg cag    548
Gly Ser Leu Tyr Met Phe Val Gln Thr Ile Leu Ser Tyr Gln Met Gln
                115                 120                 125
ccc aaa atc cat ggc aaa caa gtc ttc tgg atc aga ctg ttg ttg gtt    596
Pro Lys Ile His Gly Lys Gln Val Phe Trp Ile Arg Leu Leu Leu Val
            130                 135                 140
atc tgg tgt gga gta agt gca ctt agc atg ctg act tgc tca tca gtt    644
Ile Trp Cys Gly Val Ser Ala Leu Ser Met Leu Thr Cys Ser Ser Val
            145                 150                 155
ttg cac agt ggc aat ttt ggg act gat tta gaa cag aaa ctc cat tgg    692
Leu His Ser Gly Asn Phe Gly Thr Asp Leu Glu Gln Lys Leu His Trp
        160                 165                 170
aac ccc gag gac aaa ggt tat gcg ctt cac atg atc act act gca gca    740
Asn Pro Glu Asp Lys Gly Tyr Ala Leu His Met Ile Thr Thr Ala Ala
175                 180                 185                 190
gaa tgg tct atg tca ttt tcc ttc ttt ggt ttt ttc ctg act tac att    788
Glu Trp Ser Met Ser Phe Ser Phe Phe Gly Phe Phe Leu Thr Tyr Ile
                195                 200                 205
cgt gat ttt cag aaa att tcc tta cgg gtg gaa gcc aac tta cat gga    836
Arg Asp Phe Gln Lys Ile Ser Leu Arg Val Glu Ala Asn Leu His Gly
            210                 215                 220
tta acc ctc tat gac act gca cct tgc cct att aac aat gaa cga aca    884
Leu Thr Leu Tyr Asp Thr Ala Pro Cys Pro Ile Asn Asn Glu Arg Thr
            225                 230                 235
cgg cta ctt tcc aga gat att aga tgaaaggata aaatatttct gtaatgatta   938
Arg Leu Leu Ser Arg Asp Ile Arg
        240                 245
tgattctcag ggattgggga aaggttcaca gaagttgctt attcttctct gaaattttca   998
accacttaat caaggctgac agtaacactg atgaatgctg ataatcagga aacatgaaag  1058
aagccatttg atagattatt ctaaaggata tcatcaagaa gactattaaa aacacctatg  1118
cctatacttt tttatctcag aaataaagt caaaagacta tgaaaaaaaa aaaaaa      1174
```

<210> 144
<211> 1158
<212> DNA

156

<213> Homo sapiens
<220>
<221> polyA_signal
<222> 1133..1138
<220>
<221> polyA_site
<222> 1146..1158
<220>
<221> misc_feature
<222> 652
<223> n=a, g, c or t
<400> 144

```
aarttgagct tggggactgc agctgtgggg agatttcagt gcattgcctc ccctgggtgc      60
tcttcatctt ggatttgaaa gttgagagca gcatgttttg cccactgaaa ctcatcctgs     120
tgrsagtgta mtggattatt ccttgggcct gaatgacttg aatgtttccc cgcctgagct     180
aacagtccat gtgggtgatt cagctctg atg gga tgt gtt ttc cag agc aca        232
                                 Met Gly Cys Val Phe Gln Ser Thr
                                  1                   5
gaa gac aaa tgt ata ttc aag ata gac tgg act ctg tca cca gga gag       280
Glu Asp Lys Cys Ile Phe Lys Ile Asp Trp Thr Leu Ser Pro Gly Glu
     10                  15                  20
cac gcc aag gac gaa tat gtg cta tac tat tac tcc aat ctc agt gtg       328
His Ala Lys Asp Glu Tyr Val Leu Tyr Tyr Tyr Ser Asn Leu Ser Val
25                  30                  35                  40
cct att ggg cgc ttc cag aac cgc gta cac ttg atg ggg gac atc tta       376
Pro Ile Gly Arg Phe Gln Asn Arg Val His Leu Met Gly Asp Ile Leu
                 45                  50                  55
tgc aat gat ggc tct ctc ctg ctc caa gat gtg caa gag gct gac cag       424
Cys Asn Asp Gly Ser Leu Leu Leu Gln Asp Val Gln Glu Ala Asp Gln
                 60                  65                  70
gga acc tat atc tgt gaa atc cgc ctc aaa ggg gag agc cag gtg ttc       472
Gly Thr Tyr Ile Cys Glu Ile Arg Leu Lys Gly Glu Ser Gln Val Phe
                 75                  80                  85
aag aag gcg gtg gta ctg cat gtg ctt cca gag gag ccc aaa ggt acg       520
Lys Lys Ala Val Val Leu His Val Leu Pro Glu Glu Pro Lys Gly Thr
     90                  95                  100
caa atg ctt act taaagagggg ccaaggggca agagctttca tgtgcaagag          572
Gln Met Leu Thr
105
gcaaggaaac tgattatctt gagtaaatgc cagcctttgg gctaagtact taccacagag     632
tgaatcttca aagaaatgan tcattaaatt atttcagrtc agaataaaaa takgagttat     692
tttagttaak aataaaatat tgataattat tgtattatta ctttaaacac acttccccct     752
cacaaaagcc ctgtgaagga tgttttgttc acatataatg tccaaatatg ttttggacac     812
atatttatta aatggaataa atagtamttg aaccctggca ccthtgacaa caaagtcyat     872
gttytttta ctatgcccta atacctttsa tcagttatcc acattgatgc tacatytgta     932
ttttataggt accctatgtt aggtgttttg ggggatagaa aagaaataag cagkycaggc     992
tcagtggctc atgcctgtaa tcctagcatt ttgggaggct gaggcagcag aamtgcctga    1052
gccccagggt tcaagactgc agtgagctat gawggcacca ctgcattyta gcctgggwga    1112
cagagcaaga ctytgtttaa aataaaaaaa gagaaaaaaa aaaaaa                   1158
```

<210> 145
<211> 754
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 5..142
<223> Von Heijne matrix

157

```
        score 6.59999990463257
        seq VCCYLFWLIAILA/QL
<220>
<221> polyA_signal
<222> 716..721
<220>
<221> polyA_site
<222> 742..754
<400> 145
tgtg atg agc gtg ttc tgg ggc ttc gtc ggc ttc ttg gtg cct tgg ttc           49
     Met Ser Val Phe Trp Gly Phe Val Gly Phe Leu Val Pro Trp Phe
         -45                 -40                 -35
atc ccc aag ggt cct aac cgg gga gtt atc att acc atg ttg gtg acc            97
Ile Pro Lys Gly Pro Asn Arg Gly Val Ile Ile Thr Met Leu Val Thr
    -30                 -25                 -20
tgt tca gtt tgc tgc tat ctc ttt tgg ctg att gca att ctg gcc caa           145
Cys Ser Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala Ile Leu Ala Gln
-15                 -10                 -5                  1
ctc aac cct ctc ttt gga ccg caa ttg aaa aat gaa acc atc tgg tat           193
Leu Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu Thr Ile Trp Tyr
         5                   10                  15
ctg aag tat cat tgg cct tgaggaagaa gacatgctct acagtgctca                  241
Leu Lys Tyr His Trp Pro
         20
gtctttgagg tcacgagaag agaatgcctt ctagatgcaa aatcacctct aaaccagacc         301
acttttcttg acttgcctgt tttggccatt agctgcctta aacgttaaca gcacatttga         361
atgccttatt ctacaatgca gcgtgttttc ctttgccttt tttgcacttt ggtgaattac         421
gtgcctccat aacctgaact gtgccgactc cacaaaacga ttatgtactc ttctgagata         481
gaagatgctg ttcttctgag agatacgtta ctctctcctt ggaatctgtg gatttgaaga         541
tggctcctgc cttctcacgt gggaatcagt gaagtgttta gaaactgctg caagacaaac         601
aagactccag tggggtggtc agtaggagag cacgttcaga gggaagagcc atctcaacag         661
aatcgcacca aactatactt tcaggatgaa tttcttcttt ctgccatctt ttggaataaa         721
tattttcctc ctttctatgt aaaaaaaaaa aaa                                       754


<210> 146
<211> 1073
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 98..181
<223> Von Heijne matrix
        score 3.59999990463257
        seq PLSDSWALLPASA/GV
<220>
<221> polyA_signal
<222> 1035..1040
<220>
<221> polyA_site
<222> 1060..1073
<400> 146
ccgattacag ctaggtagtg gagcgccgct gcttacctgg gtgcaggaga cagccggagt          60
cgctggggga gctccgcgcc gccggacgcc cgtgacc atg tgg agg ctg ctg gct         115
                                         Met Trp Arg Leu Leu Ala
                                                 -25
cgc gct agt gcg ccg ctc ctg cgg gtg ccc ttg tca gat tcc tgg gca          163
Arg Ala Ser Ala Pro Leu Leu Arg Val Pro Leu Ser Asp Ser Trp Ala
    -20                 -15                 -10
```

158

```
ctc ctc ccc gcc agt gct ggc gta aag aca ctg ctc cca gta cca agt    211
Leu Leu Pro Ala Ser Ala Gly Val Lys Thr Leu Leu Pro Val Pro Ser
    -5              1               5                       10
ttt gaa gat gtt tcc att cct gaa aaa ccc aag ctt aga ttt att gaa    259
Phe Glu Asp Val Ser Ile Pro Glu Lys Pro Lys Leu Arg Phe Ile Glu
                15              20                      25
agg gca cca ctt gtg cca aaa gta aga aga gaa cct aaa aat tta agt    307
Arg Ala Pro Leu Val Pro Lys Val Arg Arg Glu Pro Lys Asn Leu Ser
            30                  35                  40
gac ata cgg gga cct tcc act gaa gct acg gag ttt aca gaa ggc aat    355
Asp Ile Arg Gly Pro Ser Thr Glu Ala Thr Glu Phe Thr Glu Gly Asn
            45                  50                  55
ttt gca atc ttg gca ttg ggt ggt ggc tac ctg cat tgg ggc cac ttt    403
Phe Ala Ile Leu Ala Leu Gly Gly Gly Tyr Leu His Trp Gly His Phe
    60                  65                  70
gaa atg atg cgc ctg aca atc aac cgc tct atg gac ccc aag aac atg    451
Glu Met Met Arg Leu Thr Ile Asn Arg Ser Met Asp Pro Lys Asn Met
75                  80                  85                  90
ttt gcc ata tgg cga gta cca gcc cct ttc aag ccc atc act cgc aaa    499
Phe Ala Ile Trp Arg Val Pro Ala Pro Phe Lys Pro Ile Thr Arg Lys
                95                  100                 105
agt gtt ggg cat cgc atg ggg gga ggc aaa ggt gct att gac cac tac    547
Ser Val Gly His Arg Met Gly Gly Gly Lys Gly Ala Ile Asp His Tyr
            110                 115                 120
gtg aca cct gtg aag gct ggc cgc ctt gtt gta gag atg ggt ggg cgt    595
Val Thr Pro Val Lys Ala Gly Arg Leu Val Val Glu Met Gly Gly Arg
            125                 130                 135
tgt gaa ttt gaa gaa gtg caa ggt ttc ctt gac cag gtt gcc cac aag    643
Cys Glu Phe Glu Glu Val Gln Gly Phe Leu Asp Gln Val Ala His Lys
    140                 145                 150
ttg ccc ttc gca gca aag gct gtg agc cgc ggg act cta gag aag atg    691
Leu Pro Phe Ala Ala Lys Ala Val Ser Arg Gly Thr Leu Glu Lys Met
155                 160                 165                 170
cga aaa gat caa gag gaa aga gaa cgt aac aac cag aac ccc tgg aca    739
Arg Lys Asp Gln Glu Glu Arg Glu Arg Asn Asn Gln Asn Pro Trp Thr
                175                 180                 185
ttt gag cga ata gcc act gcc aac atg ctg ggc ata cgg aaa gta ctg    787
Phe Glu Arg Ile Ala Thr Ala Asn Met Leu Gly Ile Arg Lys Val Leu
                190                 195                 200
agc cca tat gac ttg acc cac aag ggg aaa tac tgg ggc aag ttc tac    835
Ser Pro Tyr Asp Leu Thr His Lys Gly Lys Tyr Trp Gly Lys Phe Tyr
            205                 210                 215
atg ccc aaa cgt gtg tagtgagtgt aggagataac tgtatatagg ctactgaaag    890
Met Pro Lys Arg Val
            220
aaggattctg catttctatt cccctcagcc tacccactga agtctttggg tagctcttaa    950
gccataacta aggagcagca tttgagtaga tttctgaaaa acgatgttat ttgttgattt   1010
aaaaagaaaa ctgtattttt attaaataaa atttaaacat cacttcagga aaaaaaaaaa   1070
aaa                                                                 1073
```

<210> 147
<211> 413
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 46..189
<223> Von Heijne matrix

```
        score 4.09999990463257
        seq VFMLIVSVLALIP/ET
<220>
<221> polyA_signal
<222> 377..382
<220>
<221> polyA_site
<222> 402..413
<400> 147
tgagaagagt tgagggaaag tgctgctgct gggtctgcag acgcg atg gat aac gtg        57
                                                  Met Asp Asn Val
                                                                -45
cag ccg aaa ata aaa cat cgc ccc ttc tgc ttc agt gtg aaa ggc cac       105
Gln Pro Lys Ile Lys His Arg Pro Phe Cys Phe Ser Val Lys Gly His
            -40             -35             -30
gtg aag atg ctg cgg ctg gat att atc aac tca ctg gta aca aca gta       153
Val Lys Met Leu Arg Leu Asp Ile Ile Asn Ser Leu Val Thr Thr Val
            -25             -20             -15
ttc atg ctc atc gta tct gtg ttg gca ctg ata cca gaa acc aca aca       201
Phe Met Leu Ile Val Ser Val Leu Ala Leu Ile Pro Glu Thr Thr Thr
            -10             -5              1
ttg aca gtt ggt gga ggg gtg ttt gca ctt gtg aca gca gta tgc tgc       249
Leu Thr Val Gly Gly Gly Val Phe Ala Leu Val Thr Ala Val Cys Cys
5                   10              15              20
ctt gcc gac ggg gcc ctt att tac cgg aag ctt ctg ttc aat ccc agc       297
Leu Ala Asp Gly Ala Leu Ile Tyr Arg Lys Leu Leu Phe Asn Pro Ser
            25              30              35
ggt cct tac cag aaa aag cct gtg cat gaa aaa aaa gaa gtt ttg           342
Gly Pro Tyr Gln Lys Lys Pro Val His Glu Lys Lys Glu Val Leu
            40              45              50
taattttata ttacttttta gtttgatact aagtattaaa catatttctg tattcttcca       402
aaaaaaaaaa a                                                           413

<210> 148
<211> 609
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 139..231
<223> Von Heijne matrix
        score 4.40000009536743
        seq TCCHLGLPHPVRA/PR
<220>
<221> polyA_signal
<222> 579..584
<220>
<221> polyA_site
<222> 598..609
<400> 148
tgtcggagtt ggaaagggac gcctggtttc cccccaagcg aaccgggatg ggaagtgact        60
tcaatgagat tgaacttcag ctggattgaa agagaggcta gaagttccgc ttgccagcag       120
cctccttagt agagcgga atg agt aat acc cac acg gtg ctt gtc tca ctt        171
                     Met Ser Asn Thr His Thr Val Leu Val Ser Leu
                     -30             -25
ccc cat ccg cac ccg gcc ctc acc tgc tgt cac ctc ggc ctc cca cac       219
Pro His Pro His Pro Ala Leu Thr Cys Cys His Leu Gly Leu Pro His
-20             -15             -10             -5
```

160

```
ccg gtc cgc gct ccc cgc cct ctt cct cgc gta gaa ccg tgg gat cct        267
Pro Val Arg Ala Pro Arg Pro Leu Pro Arg Val Glu Pro Trp Asp Pro
            1               5                   10
agg tgg cag gac tca gag cta agg tat cca cag gcc atg aat tcc ttc        315
Arg Trp Gln Asp Ser Glu Leu Arg Tyr Pro Gln Ala Met Asn Ser Phe
        15                  20                  25
cta aat gag cgg tca tcg ccg tgc agg acc tta agg caa gaa gca tcg        363
Leu Asn Glu Arg Ser Ser Pro Cys Arg Thr Leu Arg Gln Glu Ala Ser
    30                  35                  40
gct gac aga tgt gat ctc tgaacctgat agattgctga ttttatctta              411
Ala Asp Arg Cys Asp Leu
45                  50
ttttatcctt gacttggtac aagttttggg atttctgaaa agaccataca gataaccaca      471
aatatcaaga aagtcgtctt cagtattaag tagaatttag atttaggttt ccttcctgct      531
tcccacctcc ttcgaataag gaaacgtctt tgggaccaac tttatggaat aaataagctg      591
agctgcaaaa aaaaaaaa                                                    609


<210> 149
<211> 522
<212> DNA
<213> Homo sapiens
<220>
<221> polyA_site
<222> 512..522
<220>
<221> misc_feature
<222> 11
<223> n=a, g, c or t
<400> 149
ccaactgcag nttcgaattt accgagcgga gaggagatgc acacggcact cgagtgtgag        60
gaaaaataga a atg aag gta cat atg cac aca aaa ttt tgc ctc att tgt       110
            Met Lys Val His Met His Thr Lys Phe Cys Leu Ile Cys
              1               5                   10
ttg ctg aca ttt att ttt cat cat tgc aac cat tgc cat gaa gaa cat       158
Leu Leu Thr Phe Ile Phe His His Cys Asn His Cys His Glu Glu His
    15                  20                  25
gac cat ggc cct gaa gcg ctt cac aga cag cat cgt gga atg aca gaa       206
Asp His Gly Pro Glu Ala Leu His Arg Gln His Arg Gly Met Thr Glu
30                  35                  40                  45
ttg gag cca agc aaa ttt tca aag caa gct gct gaa aat gaa aaa aaa       254
Leu Glu Pro Ser Lys Phe Ser Lys Gln Ala Ala Glu Asn Glu Lys Lys
                50                  55                  60
tac tat att gaa aaa ctt ttt gag cgt tat ggt gaa aat gga aga tta       302
Tyr Tyr Ile Glu Lys Leu Phe Glu Arg Tyr Gly Glu Asn Gly Arg Leu
            65                  70                  75
tcc ttt ttt ggt ttg gag aaa ctt tta aca aac ttg ggc ctt gga gag       350
Ser Phe Phe Gly Leu Glu Lys Leu Leu Thr Asn Leu Gly Leu Gly Glu
            80                  85                  90
aga aaa gta gtt gag att aat cat gag gat ctt ggc cac gat cat gtt       398
Arg Lys Val Val Glu Ile Asn His Glu Asp Leu Gly His Asp His Val
    95                  100                 105
tct cat tta ggt att ttg gca gtt caa gag gga aag cat ttt cac tca       446
Ser His Leu Gly Ile Leu Ala Val Gln Glu Gly Lys His Phe His Ser
110                 115                 120                 125
cat aac cac cag cat tcc cat aat cat tta aat tca gaa aat caa act       494
His Asn His Gln His Ser His Asn His Leu Asn Ser Glu Asn Gln Thr
                130                 135                 140
gtg acc agt gta tcc aca aaaaaaaaaa                                     522
```

161

Val Thr Ser Val Ser Thr
              145

<210> 150
<211> 1322
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 126..260
<223> Von Heijne matrix
      score 4.59999990463257
      seq VLVYLVTAERVWS/DD
<220>
<221> polyA_signal
<222> 1283..1288
<220>
<221> polyA_site
<222> 1309..1322
<400> 150
ccgaaaacct tccccgcttc tggatatgaa attcaagctg cttgctgagt cctattgccg   60
gctgctggga gccaggagag ccctgaggag tagtcactca gtagcagctg acgcgtgggt   120
ccacc atg aac tgg agt atc ttt gag gga ctc ctg agt ggg gtc aac aag   170
      Met Asn Trp Ser Ile Phe Glu Gly Leu Leu Ser Gly Val Asn Lys
      -45   -40   -35
tac tcc aca gcc ttt ggg cgc atc tgg ctg tct ctg gtc ttc atc ttc   218
Tyr Ser Thr Ala Phe Gly Arg Ile Trp Leu Ser Leu Val Phe Ile Phe
-30   -25   -20   -15
cgc gtg ctg gtg tac ctg gtg acg gcc gag cgt gtg tgg agt gat gac   266
Arg Val Leu Val Tyr Leu Val Thr Ala Glu Arg Val Trp Ser Asp Asp
                  -10   -5   1
cac aag gac ttc gac tgc aat act cgc cag ccc ggc tgc tcc aac gtc   314
His Lys Asp Phe Asp Cys Asn Thr Arg Gln Pro Gly Cys Ser Asn Val
          5   10   15
tgc ttt gat gag ttc ttc cct gtg tcc cat gtg cgc ctc tgg gcc ctg   362
Cys Phe Asp Glu Phe Phe Pro Val Ser His Val Arg Leu Trp Ala Leu
    20   25   30
cag ctt atc ctg gtg aca tgc ccc tca ctg ctc gtg gtc atg cac gtg   410
Gln Leu Ile Leu Val Thr Cys Pro Ser Leu Leu Val Val Met His Val
35   40   45   50
gcc tac cgg gag gtt cag gag aag agg cac cga gaa gcc cat ggg gag   458
Ala Tyr Arg Glu Val Gln Glu Lys Arg His Arg Glu Ala His Gly Glu
                  55   60   65
aac agt ggg cgc ctc tac ctg aac ccc ggc aag aag cgg ggt ggg ctc   506
Asn Ser Gly Arg Leu Tyr Leu Asn Pro Gly Lys Lys Arg Gly Gly Leu
                  70   75   80
tgg tgg aca tat gtc tgc agc cta gtg ttc aag gcg agc gtg gac atc   554
Trp Trp Thr Tyr Val Cys Ser Leu Val Phe Lys Ala Ser Val Asp Ile
          85   90   95
gcc ttt ctc tat gtg ttc cac tca ttc tac ccc aaa tat atc ctc cct   602
Ala Phe Leu Tyr Val Phe His Ser Phe Tyr Pro Lys Tyr Ile Leu Pro
    100   105   110
cct gtg gtc aag tgc cac gca gat cca tgt ccc aat ata gtg gac tgc   650
Pro Val Val Lys Cys His Ala Asp Pro Cys Pro Asn Ile Val Asp Cys
115   120   125   130
ttc atc tcc aag ccc tca gag aag aac att ttc acc ctc ttc atg gtg   698
Phe Ile Ser Lys Pro Ser Glu Lys Asn Ile Phe Thr Leu Phe Met Val
                  135   140   145

162

```
gcc aca gct gcc atc tgc atc ctg ctc aac ctc gtg gag ctc atc tac        746
Ala Thr Ala Ala Ile Cys Ile Leu Leu Asn Leu Val Glu Leu Ile Tyr
        150                     155                 160
ctg gtg agc aag aga tgc cac gag tgc ctg gca gca agg aaa gct caa        794
Leu Val Ser Lys Arg Cys His Glu Cys Leu Ala Ala Arg Lys Ala Gln
        165                     170                 175
gcc atg tgc aca ggt cat cac ccc cac gat acc acc tct tcc tgc aaa        842
Ala Met Cys Thr Gly His His Pro His Asp Thr Thr Ser Ser Cys Lys
    180                     185                 190
caa gac gac ctc ctt tcg ggt gac ctc atc ttt ctg ggc tca gac agt        890
Gln Asp Asp Leu Leu Ser Gly Asp Leu Ile Phe Leu Gly Ser Asp Ser
195                     200                 205                 210
cat cct cct ctc tta cca gac cgc ccc cga gac cat gtg aag aaa acc        938
His Pro Pro Leu Leu Pro Asp Arg Pro Arg Asp His Val Lys Lys Thr
                215                 220                 225
atc ttg tgaggggctg cctggactgg tctggcaggt tgggcctgga tggggaggct        994
Ile Leu
ctagcatctc tcataggtgc aacctgagag tgggggagct aagccatgag gtaggggcag       1054
gcaagagaga ggattcagac gctctgggag ccagttccta gtcctcaact ccagccacct       1114
gccccagctc gacggcactg gccagttcc cctctgctc tgcagctcgg tttccttttc        1174
tagaatggaa atagtgaggg ccaatgccca gggttggagg gaggagggcg ttcatagaag       1234
aacacacatg cgggcacctt catcgtgtgt ggcccactgt cagaacttaa taaaagtcaa       1294
ctcatttgct ggttaaaaaa aaaaaaaaa                                       1322


<210> 151
<211> 1290
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 50..160
<223> Von Heijne matrix
      score 4
      seq PLSLDCGHSLCRA/CI
<220>
<221> polyA_site
<222> 1280..1290
<400> 151
gaggagagcc tcaggagtta ggaccagaag aagccaggga agcagtgca atg gct tca        58
                                                        Met Ala Ser
                                                                -35
aaa atc ttg ctt aac gta caa gag gag gtg acc tgt ccc atc tgc ctg       106
Lys Ile Leu Leu Asn Val Gln Glu Glu Val Thr Cys Pro Ile Cys Leu
            -30                 -25                 -20
gag ctg ttg aca gaa ccc ttg agt cta gac tgt ggc cac agc ctc tgc       154
Glu Leu Leu Thr Glu Pro Leu Ser Leu Asp Cys Gly His Ser Leu Cys
            -15                 -10                 -5
cga gcc tgc atc act gtg agc aac aag gag gca gtg acc agc atg gga       202
Arg Ala Cys Ile Thr Val Ser Asn Lys Glu Ala Val Thr Ser Met Gly
    1                   5                   10
gga aaa agc agc tgt cct gtg tgt ggt atc agt tac tca ttt gaa cat       250
Gly Lys Ser Ser Cys Pro Val Cys Gly Ile Ser Tyr Ser Phe Glu His
15                  20                  25                  30
cta cag gct aat cag cat ctg gcc aac ata gtg gag aga ctc aag gag       298
Leu Gln Ala Asn Gln His Leu Ala Asn Ile Val Glu Arg Leu Lys Glu
            35                  40                  45
gtc aag ttg agc cca gac aat ggg aag aag aga gat ctc tgt gat cat       346
Val Lys Leu Ser Pro Asp Asn Gly Lys Lys Arg Asp Leu Cys Asp His
```

```
                50                        55                        60
cat gga gag aaa ctc cta ctc ttc tgt aag gag gat agg aaa gtc att      394
His Gly Glu Lys Leu Leu Leu Phe Cys Lys Glu Asp Arg Lys Val Ile
            65                        70                        75
tgc tgg ctt tgt gag cgg tct cag gag cac cgt ggt cac cac aca gtc      442
Cys Trp Leu Cys Glu Arg Ser Gln Glu His Arg Gly His His Thr Val
        80                        85                        90
ctc acg gag gaa gta ttc aag gaa tgt cag gag aaa ctc cag gca gtc      490
Leu Thr Glu Glu Val Phe Lys Glu Cys Gln Glu Lys Leu Gln Ala Val
95                        100                       105                       110
ctc aag agg ctg aag aag gaa gag gag gaa gct gag aag ctg gaa gct      538
Leu Lys Arg Leu Lys Lys Glu Glu Glu Glu Ala Glu Lys Leu Glu Ala
                    115                       120                       125
gac atc aga gaa gag aaa act tcc tgg aag tat cag gta caa act gag      586
Asp Ile Arg Glu Glu Lys Thr Ser Trp Lys Tyr Gln Val Gln Thr Glu
            130                       135                       140
aga caa agg ata caa aca gaa ttt gat cag ctt aga agc atc cta aat      634
Arg Gln Arg Ile Gln Thr Glu Phe Asp Gln Leu Arg Ser Ile Leu Asn
        145                       150                       155
aat gag gag cag aga gag ctg caa aga ttg gaa gaa gaa gaa aag aag      682
Asn Glu Glu Gln Arg Glu Leu Gln Arg Leu Glu Glu Glu Glu Lys Lys
    160                       165                       170
acg ctg gat aag ttt gca gag gct gag gat gag cta gtt cag cag aag      730
Thr Leu Asp Lys Phe Ala Glu Ala Glu Asp Glu Leu Val Gln Gln Lys
175                       180                       185                       190
cag ttg gtg aga gag ctc atc tca gat gtg gag tgt cgg agt cag tgg      778
Gln Leu Val Arg Glu Leu Ile Ser Asp Val Glu Cys Arg Ser Gln Trp
                    195                       200                       205
tca aca atg gag ctg ctg cag gac atg agt gga atc atg aaa tgg agt      826
Ser Thr Met Glu Leu Leu Gln Asp Met Ser Gly Ile Met Lys Trp Ser
                210                       215                       220
gag atc tgg agg ctg aaa aag cca aaa atg gtt tcc aag aaa ctg aag      874
Glu Ile Trp Arg Leu Lys Lys Pro Lys Met Val Ser Lys Lys Leu Lys
            225                       230                       235
act gta ttc cat gct cca gat ctg agt agg atg ctg caa atg ttt aga      922
Thr Val Phe His Ala Pro Asp Leu Ser Arg Met Leu Gln Met Phe Arg
        240                       245                       250
gaa ctg aca gct gtc cgg tgc tac tgg gtg gat gtc aca ctg aat tca      970
Glu Leu Thr Ala Val Arg Cys Tyr Trp Val Asp Val Thr Leu Asn Ser
255                       260                       265                       270
gtc aac cta aat ttg aat ctt gtc ctt tca gaa gat cag aga caa gtg      1018
Val Asn Leu Asn Leu Asn Leu Val Leu Ser Glu Asp Gln Arg Gln Val
                    275                       280                       285
ata tct gtg cca att tgg cct ttt cag tgt tat aat tat ggt gtc ttg      1066
Ile Ser Val Pro Ile Trp Pro Phe Gln Cys Tyr Asn Tyr Gly Val Leu
                290                       295                       300
gga tcc caa tat ttc tcc tct ggg aaa cat tac tgg gaa gtg gac gtg      1114
Gly Ser Gln Tyr Phe Ser Ser Gly Lys His Tyr Trp Glu Val Asp Val
            305                       310                       315
tcc aag aaa act gcc tgg atc ctg ggg gta tac tgt aga aca tat tcc      1162
Ser Lys Lys Thr Ala Trp Ile Leu Gly Val Tyr Cys Arg Thr Tyr Ser
        320                       325                       330
cgc cat atg aag tat gtt gtt aga aga tgt gca aat cgt caa aat ctt      1210
Arg His Met Lys Tyr Val Val Arg Arg Cys Ala Asn Arg Gln Asn Leu
335                       340                       345                       350
tac acc aaa tac aga cct cta ttt ggc tac tgg gtt ata ggg tta cag      1258
Tyr Thr Lys Tyr Arg Pro Leu Phe Gly Tyr Trp Val Ile Gly Leu Gln
                    355                       360                       365
```

164

```
aat aaa tgt aag tat ggt gcc aaaaaaaaaa a                                    1290
Asn Lys Cys Lys Tyr Gly Ala
            370


<210> 152
<211> 1364
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 83..139
<223> Von Heijne matrix
      score 8.60000038146973
      seq LLWLALACSPVHT/TL
<220>
<221> polyA_site
<222> 1356..1354
<400> 152
gcctgggagc tgaggcagcc accgtctcag cctggccagc cctctggacc ccgaggttgg         60
accctactgt gacacaccta cc atg cgg aca ctc ttc aac ctc ctc tgg ctt         112
                          Met Arg Thr Leu Phe Asn Leu Leu Trp Leu
                              -15                         -10
gcc ctg gcc tgc agc cct gtt cac act acc ctg tca aag tca gat gcc         160
Ala Leu Ala Cys Ser Pro Val His Thr Thr Leu Ser Lys Ser Asp Ala
            -5                      1               5
aaa aaa gcc gcc tca aag acg ctg ctg gag aag agt cag ttt tca gat         208
Lys Lys Ala Ala Ser Lys Thr Leu Leu Glu Lys Ser Gln Phe Ser Asp
        10                  15                  20
aag ccg gtg caa gac cgg ggt ttg gtg gtg acg gac ctc aaa gct gag         256
Lys Pro Val Gln Asp Arg Gly Leu Val Val Thr Asp Leu Lys Ala Glu
        25                  30                  35
agt gtg gtt ctt gag cat cgc agc tac tgc tcg gca aag gcc cgg gac         304
Ser Val Val Leu Glu His Arg Ser Tyr Cys Ser Ala Lys Ala Arg Asp
40                  45                  50                  55
aga cac ttt gct ggg gat gta ctg ggc tat gtc act cca tgg aac agc         352
Arg His Phe Ala Gly Asp Val Leu Gly Tyr Val Thr Pro Trp Asn Ser
                60                  65                  70
cat ggc tac gat gtc acc aag gtc ttt ggg agc aag ttc aca cag atc         400
His Gly Tyr Asp Val Thr Lys Val Phe Gly Ser Lys Phe Thr Gln Ile
                75                  80                  85
tca ccc gtc tgg ctg cag ttg aag aga cgt ggc cgt gag atg ttt gag         448
Ser Pro Val Trp Leu Gln Leu Lys Arg Arg Gly Arg Glu Met Phe Glu
                90                  95                  100
gtc acg ggc ctc cac gac gtg gac caa ggg tgg atg cga gct gtc agg         496
Val Thr Gly Leu His Asp Val Asp Gln Gly Trp Met Arg Ala Val Arg
                105                 110                 115
aag cat gcc aag ggc ctg cac ata gtg cct cgg ctc ctg ttt gag gac         544
Lys His Ala Lys Gly Leu His Ile Val Pro Arg Leu Leu Phe Glu Asp
120                 125                 130                 135
tgg act tac gat gat ttc cgg aac gtc tta gac agt gag gat gag ata         592
Trp Thr Tyr Asp Asp Phe Arg Asn Val Leu Asp Ser Glu Asp Glu Ile
                140                 145                 150
gag gag ctg agc aag acc gtg gtc cag gtg gca aag aac cag cat ttc         640
Glu Glu Leu Ser Lys Thr Val Val Gln Val Ala Lys Asn Gln His Phe
                155                 160                 165
gat ggc ttc gtg gtg gag gtc tgg aac cag ctg cta agc cag aag cgc         688
Asp Gly Phe Val Val Glu Val Trp Asn Gln Leu Leu Ser Gln Lys Arg
                170                 175                 180
```

165

```
gtg ggc ctc atc cac atg ctc acc cac ttg gcc gag gcc ctg cac cag        736
Val Gly Leu Ile His Met Leu Thr His Leu Ala Glu Ala Leu His Gln
        185                 190                 195
gcc cgg ctg ctg gcc ctc ctg gtc atc ccg cct gcc atc acc ccc ggg        784
Ala Arg Leu Leu Ala Leu Leu Val Ile Pro Pro Ala Ile Thr Pro Gly
200                 205                 210                 215
acc gac cag ctg ggc atg ttc acg cac aag gag ttt gag cag ctg gcc        832
Thr Asp Gln Leu Gly Met Phe Thr His Lys Glu Phe Glu Gln Leu Ala
                220                 225                 230
ccc gtg ctg gat ggt ttc agc ctc atg acc tac gac tac tct aca gcg        880
Pro Val Leu Asp Gly Phe Ser Leu Met Thr Tyr Asp Tyr Ser Thr Ala
                235                 240                 245
cat cag cct ggc cct aat gca ccc ctg tcc tgg gtt cga gcc tgc gtc        928
His Gln Pro Gly Pro Asn Ala Pro Leu Ser Trp Val Arg Ala Cys Val
                250                 255                 260
cag gtc ctg gac ccg aag tcc aag tgg cga agc aaa atc ctc ctg ggg        976
Gln Val Leu Asp Pro Lys Ser Lys Trp Arg Ser Lys Ile Leu Leu Gly
                265                 270                 275
ctc aac ttc tat ggt atg gac tac gcg acc tcc aag gat gcc cgt gag       1024
Leu Asn Phe Tyr Gly Met Asp Tyr Ala Thr Ser Lys Asp Ala Arg Glu
280                 285                 290                 295
cct gtt gtc ggg gcc agg tac atc cag aca ctg aag gac cac agg ccc       1072
Pro Val Val Gly Ala Arg Tyr Ile Gln Thr Leu Lys Asp His Arg Pro
                300                 305                 310
cgg atg gtg tgg gac agc cag gcc tca gag cac ttc ttc gag tac aag       1120
Arg Met Val Trp Asp Ser Gln Ala Ser Glu His Phe Phe Glu Tyr Lys
                315                 320                 325
aag agc cgc agt ggg agg cac gtc gtc ttc tac cca acc ctg aag tcc       1168
Lys Ser Arg Ser Gly Arg His Val Val Phe Tyr Pro Thr Leu Lys Ser
                330                 335                 340
ctg cag gtg cgg ctg gag ctg gcc cgg gag ctg ggc gtt ggg gtc tct       1216
Leu Gln Val Arg Leu Glu Leu Ala Arg Glu Leu Gly Val Gly Val Ser
                345                 350                 355
atc tgg gag ctg ggc cag ggc ctg gac tac ttc tac gac ctg ctc           1261
Ile Trp Glu Leu Gly Gln Gly Leu Asp Tyr Phe Tyr Asp Leu Leu
360                 365                 370
taggtgggca ttgcggcctc cgcggtggac gtgttctttt ctaagccatg gagtgagtga     1321
gcaggtgtga aatacaggcc tccactccgt ttgcaaaaaa aaa                        1364
```

<210> 153
<211> 1470
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 57..95
<223> Von Heijne matrix
. score 3.90000009536743
    seq MLLSIGMLMLSAT/QV
<220>
<221> polyA_signal
<222> 1438..1443
<220>
<221> polyA_site
<222> 1458..1470
<400> 153

```
gctggcaaga ctgtttgtgt tgcgggggcc ggacttcaag gtgattttac aacgag atg     59
                                                                    Met
```

```
ctg ctc tcc ata ggg atg ctc atg ctg tca gcc aca caa gtc tac acc      107
Leu Leu Ser Ile Gly Met Leu Met Leu Ser Ala Thr Gln Val Tyr Thr
        -10             -5                  1
gtc ttg act gtc cag ctc ttt gca ttc tta aac cca ctg cct gta gaa      155
Val Leu Thr Val Gln Leu Phe Ala Phe Leu Asn Pro Leu Pro Val Glu
5               10                  15                  20
gca gac att tta gca tat aac ttt gaa aat gca tct cag aca ttt gat      203
Ala Asp Ile Leu Ala Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe Asp
            25                  30                  35
gac ctc cct gca aga ttt ggt tat aga ctt cca gct gaa ggt tta aag      251
Asp Leu Pro Ala Arg Phe Gly Tyr Arg Leu Pro Ala Glu Gly Leu Lys
            40                  45                  50
ggt ttt tta att aac tca aaa cca gag aat gcc tgt gaa ccc ata gtg      299
Gly Phe Leu Ile Asn Ser Lys Pro Glu Asn Ala Cys Glu Pro Ile Val
        55                  60                  65
cct cca cca gta aaa gac aat tca tct ggc act ttc atc gtg tta att      347
Pro Pro Pro Val Lys Asp Asn Ser Ser Gly Thr Phe Ile Val Leu Ile
        70                  75                  80
aga aga ctt gat tgt aat ttt gat ata aag gtt tta aat gca cag aga      395
Arg Arg Leu Asp Cys Asn Phe Asp Ile Lys Val Leu Asn Ala Gln Arg
85                  90                  95                  100
gca gga tac aag gca gcc ata gtt cac aat gtt gat tct gat gac ctc      443
Ala Gly Tyr Lys Ala Ala Ile Val His Asn Val Asp Ser Asp Asp Leu
                105                 110                 115
att agc atg gga tcc aac gac att gag gta cta aag aaa att gac att      491
Ile Ser Met Gly Ser Asn Asp Ile Glu Val Leu Lys Lys Ile Asp Ile
                120                 125                 130
cca tct gtc ttt att ggt gaa tca tca gct agt tct ctg aaa gat gaa      539
Pro Ser Val Phe Ile Gly Glu Ser Ser Ala Ser Ser Leu Lys Asp Glu
                135                 140                 145
ttc aca tat gaa aaa ggg ggc cac ctt atc tta gtt cca gaa ttt agt      587
Phe Thr Tyr Glu Lys Gly Gly His Leu Ile Leu Val Pro Glu Phe Ser
        150                 155                 160
ctt cct ttg gaa tac tac cta att ccc ttc ctt atc ata gtg ggc atc      635
Leu Pro Leu Glu Tyr Tyr Leu Ile Pro Phe Leu Ile Ile Val Gly Ile
165                 170                 175                 180
tgt ctc atc ttg ata gtc att ttc atg atc aca aaa ttt gtc cag gat      683
Cys Leu Ile Leu Ile Val Ile Phe Met Ile Thr Lys Phe Val Gln Asp
                185                 190                 195
aga cat aga gct aga aga aac aga ctt cgt aaa gat caa ctt aag aaa      731
Arg His Arg Ala Arg Arg Asn Arg Leu Arg Lys Asp Gln Leu Lys Lys
                200                 205                 210
ctt cct gta cat aaa ttc aag aaa gga gat gag tat gat gta tgt gcc      779
Leu Pro Val His Lys Phe Lys Lys Gly Asp Glu Tyr Asp Val Cys Ala
        215                 220                 225
att tgt ttg gat gag tat gaa gat gga gac aaa ctc aga atc ctt ccc      827
Ile Cys Leu Asp Glu Tyr Glu Asp Gly Asp Lys Leu Arg Ile Leu Pro
        230                 235                 240
tgt tcc cat gct tat cat tgc aag tgt gta gac cct tgg cta act aaa      875
Cys Ser His Ala Tyr His Cys Lys Cys Val Asp Pro Trp Leu Thr Lys
245                 250                 255                 260
acc aaa aaa acc tgt cca gtg tgc agg caa aaa gtt gtt cct tct caa      923
Thr Lys Lys Thr Cys Pro Val Cys Arg Gln Lys Val Val Pro Ser Gln
                265                 270                 275
ggc gat tca gac tct gac aca gac agt agt caa gaa gaa aat gaa gtg      971
Gly Asp Ser Asp Ser Asp Thr Asp Ser Ser Gln Glu Glu Asn Glu Val
                280                 285                 290
aca gaa cat acc cct tta ctg aga cct tta gct tct gtc agt gcc cag     1019
```

296

```
           Thr Glu His Thr Pro Leu Leu Arg Pro Leu Ala Ser Val Ser Ala Gln
                   295                 300                 305
tca ttt ggg gct tta tcg gaa tcc cgc tca cat cag aac atg aca gaa          1067
Ser Phe Gly Ala Leu Ser Glu Ser Arg Ser His Gln Asn Met Thr Glu
        310                 315                 320
tct tca gac tat gag gaa gac gac aat gaa gat act gac agt agt gat          1115
Ser Ser Asp Tyr Glu Glu Asp Asp Asn Glu Asp Thr Asp Ser Ser Asp
325                 330                 335                 340
gca gaa aat gaa att aat gaa cat gat gtc gtg gtc cag ttg cag cct          1163
Ala Glu Asn Glu Ile Asn Glu His Asp Val Val Val Gln Leu Gln Pro
                345                 350                 355
aat ggt gaa cgg gat tac aac ata gca aat act gtt tgactttcag             1209
Asn Gly Glu Arg Asp Tyr Asn Ile Ala Asn Thr Val
                360                 365
aagatgattg gtttatttcc ctttaaaatg attaggtata tactgtaatt tgattttttg       1269
ctcccttaaa agatttctgt agaaataact tattttttag tactctacag tttaatcaaa       1329
ttactgaaac aggacttttg atctggtatt tatctgccaa gaatatactt cattcactaa       1389
taatagactg gtgctgtaac tcaagcatca attcagctct tcttttggaa tgaaagtata       1449
gccaaaacaa aaaaaaaaaa a                                                 1470


<210> 154
<211> 982
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 72..197
<223> Von Heijne matrix
        score 7.19999980926514
        seq ILFSLSFLLVIIT/FP
<220>
<221> polyA_site
<222> 970..982
<400> 154
gctgcctgtt cttcacactt agctccaaac ccatgaaaaa ttgccaagta taaaagcttc       60
tcaagaatga g atg gat tct agg gtg tct tca cct gag aag caa gat aaa        110
            Met Asp Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys
                -40                 -35                 -30
gag aat ttc gtg ggt gtc aac aat aaa cgg ctt ggt gta tgt ggc tgg        158
Glu Asn Phe Val Gly Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp
        -25                 -20                 -15
atc ctg ttt tcc ctc tct ttc ctg ttg gtg atc att acc ttc ccc atc        206
Ile Leu Phe Ser Leu Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile
        -10                 -5                  1
tcc ata tgg atg tgc ttg aag atc att agg gag tat gaa cgt gct gtt        254
Ser Ile Trp Met Cys Leu Lys Ile Ile Arg Glu Tyr Glu Arg Ala Val
5                   10                  15
gta ttc cgt ctg gga cgc atc caa gct gac aaa gcc aag ggg cca ggt        302
Val Phe Arg Leu Gly Arg Ile Gln Ala Asp Lys Ala Lys Gly Pro Gly
20                  25                  30                  35
ttg atc ctg gtc ctg cca tgc ata gat gtg ttt gtc aag gtt gac ctc        350
Leu Ile Leu Val Leu Pro Cys Ile Asp Val Phe Val Lys Val Asp Leu
                40                  45                  50
cga aca gtt act tgc aac att cct cca caa gag atc ctc acc aga gac        398
Arg Thr Val Thr Cys Asn Ile Pro Pro Gln Glu Ile Leu Thr Arg Asp
                55                  60                  65
tcc gta act act cag gta gat gga gtt gtc tat tac aga atc tat agt        446
Ser Val Thr Thr Gln Val Asp Gly Val Val Tyr Tyr Arg Ile Tyr Ser
```

```
         70                 75                 80
gct gtc tca gca gtg gct aat gtc aac gat gtc cat caa gca aca ttt     494
Ala Val Ser Ala Val Ala Asn Val Asn Asp Val His Gln Ala Thr Phe
     85                 90                 95
ctg ctg gct caa acc act ctg aga aat gtc tta ggg aca cag acc ttg     542
Leu Leu Ala Gln Thr Thr Leu Arg Asn Val Leu Gly Thr Gln Thr Leu
100                105                110                115
tcc cag atc tta gct gga cga gaa gag atc gcc cat agc atc cag act     590
Ser Gln Ile Leu Ala Gly Arg Glu Glu Ile Ala His Ser Ile Gln Thr
                   120                125                130
tta ctt gat gat gcc acc gaa ctg tgg ggg atc cgg gtg gcc cga gtg     638
Leu Leu Asp Asp Ala Thr Glu Leu Trp Gly Ile Arg Val Ala Arg Val
                   135                140                145
gaa atc aaa gat gtt cgg att ccc gtg cag ttg cag aga tcc atg gca     686
Glu Ile Lys Asp Val Arg Ile Pro Val Gln Leu Gln Arg Ser Met Ala
                   150                155                160
gcc gag gct gag gcc acc cgg gaa gcg aga gcc aag gtc ctt gca gct     734
Ala Glu Ala Glu Ala Thr Arg Glu Ala Arg Ala Lys Val Leu Ala Ala
                   165                170                175
gaa gga gaa atg agt gct tcc aaa tcc ctg aag tca gcc tcc atg gtg     782
Glu Gly Glu Met Ser Ala Ser Lys Ser Leu Lys Ser Ala Ser Met Val
180                185                190                195
ctg gct gag tct ccc ata gct ctc cag ctg cgc tac ctg cag acc ttg     830
Leu Ala Glu Ser Pro Ile Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu
                   200                205                210
agc acg gta gcc acc gag aag aat tct acg att gtg ttt cct ctg ccc     878
Ser Thr Val Ala Thr Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro
                   215                220                225
atg aat ata cta gag ggc att ggt ggc gtc agc tat gat aac cac aag     926
Met Asn Ile Leu Glu Gly Ile Gly Gly Val Ser Tyr Asp Asn His Lys
                   230                235                240
aag ctt cca aat aaa gcc tgaggtcctc ttgcggtagt cagctaaaaa aaaaaaaa    982
Lys Leu Pro Asn Lys Ala
                   245


<210> 155
<211> 455
<212> DNA
<213> Homo sapiens
<220>
<221> polyA_signal
<222> 425..430
<220>
<221> polyA_site
<222> 443..455
<400> 155
gtt atg cca ccc aga aac cta ctg gag tta ctt att aac atc aag gct      48
    Met Pro Pro Arg Asn Leu Leu Glu Leu Leu Ile Asn Ile Lys Ala
    1               5                 10                 15
gga acc tat ttg cct cag tcc tat ctg att cat gag cac atg gtt att      96
Gly Thr Tyr Leu Pro Gln Ser Tyr Leu Ile His Glu His Met Val Ile
                   20                 25                 30
act gat cgc atc gaa aac att gat cac ctg ggt ttc ttt att tat cga     144
Thr Asp Arg Ile Glu Asn Ile Asp His Leu Gly Phe Phe Ile Tyr Arg
                   35                 40                 45
ctg tgt cat gac aag gaa act tac aaa ctg caa cgc aga gaa act att     192
Leu Cys His Asp Lys Glu Thr Tyr Lys Leu Gln Arg Arg Glu Thr Ile
                   50                 55                 60
```

169

```
aaa ggt att cag aaa cgt gaa gcc agc aat tgt ttc gca att cgg cat        240
Lys Gly Ile Gln Lys Arg Glu Ala Ser Asn Cys Phe Ala Ile Arg His
      65                  70                  75
ttt gaa aac aaa ttt gcc gtg gaa act tta att tgt tct tgaacagtca        289
Phe Glu Asn Lys Phe Ala Val Glu Thr Leu Ile Cys Ser
80                  85                  90
agaaaaacat tattgaggaa aattaatatc acagcataac cccacccttt acattttgtg     349
cagtgattat tttttaaagt cttctttcat gtaagtagca aacagggctt tactatcttt     409
tcatctcatt aattcaatta aaaccattac cccaaaaaaa aaaaaa                    455


<210> 156
<211> 738
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 90..278
<223> Von Heijne matrix
      score 3.5
      seq GLVCAGLADMARP/AE
<220>
<221> polyA_signal
<222> 704..709
<220>
<221> polyA_site
<222> 724..738
<400> 156
gggaaaagtg actagctccc cttcgttgtc agccagggac gagaacacag ccacgctccc      60
acccggctgc aacgatccc tcggcggcg atg tcg gcc gcc ggt gcc cga ggc         113
                                Met Ser Ala Ala Gly Ala Arg Gly
                                              -60
ctg cgg gcc acc tac cac cgg ctc ccc gat aaa gtg gag ctg atg ctg        161
Leu Arg Ala Thr Tyr His Arg Leu Pro Asp Lys Val Glu Leu Met Leu
-55                  -50                  -45                  -40
ccc gag aaa ttg agg ccg ttg tac aac cat cca gca ggt ccc aga aca        209
Pro Glu Lys Leu Arg Pro Leu Tyr Asn His Pro Ala Gly Pro Arg Thr
              -35                  -30                  -25
gtt ttc ttc tgg gct cca att atg aaa tgg ggg ttg gtg tgt gct gga        257
Val Phe Phe Trp Ala Pro Ile Met Lys Trp Gly Leu Val Cys Ala Gly
              -20                  -15                  -10
ttg gct gat atg gcc aga cct gca gaa aaa ctt agc aca gct caa tct        305
Leu Ala Asp Met Ala Arg Pro Ala Glu Lys Leu Ser Thr Ala Gln Ser
              -5                   1                    5
gct gtt ttg atg gct aca ggg ttt att tgg tca aga tac tca ctt gta        353
Ala Val Leu Met Ala Thr Gly Phe Ile Trp Ser Arg Tyr Ser Leu Val
10                  15                   20                   25
att att ccg aaa aat tgg agt ctg ttt gct gtt aat ttc ttt gtg ggg        401
Ile Ile Pro Lys Asn Trp Ser Leu Phe Ala Val Asn Phe Phe Val Gly
                    30                   35                   40
gca gca gga gcc tct cag ctt ttt cgt att tgg aga tat aac caa gaa        449
Ala Ala Gly Ala Ser Gln Leu Phe Arg Ile Trp Arg Tyr Asn Gln Glu
              45                   50                   55
cta aaa gct aaa gca cac aaa taaaagagtt cctgatcacc tgaacaatct           500
Leu Lys Ala Lys Ala His Lys
              60
agatgtggac aaaaccattg ggacctagtt tattatttgg ttattgataa agcaaagcta     560
actgtgtgtt tagaaggcac tgtaactggt agctagttct tgattcaata gaaaaatgca     620
gcaaactttt aataacagtc tctctacatg acttaaggaa cttatctatg gatattagta     680
```

170 299

acattttct accatttgtc cgtaataaac catacttgct cgtaaaaaaa aaaaaaaa   738

<210> 157
<211> 649
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 88..147
<223> Von Heijne matrix
      score 12.3999996185303
      seq ALLLGALLGTAWA/RR
<220>
<221> polyA_signal
<222> 619..624
<220>
<221> polyA_site
<222> 637..649
<400> 157
ccaaagtgag agtccagcgg tcttccagcg cttgggccac ggcggcggcc ctgggagcag   60
aggaggagcg accccattac gctaaag atg aaa ggc tgg ggt tgg ctg gcc ctg   114
                              Met Lys Gly Trp Gly Trp Leu Ala Leu
                              -20   -15
ctt ctg ggg gcc ctg ctg gga acc gcc tgg gct cgg agg agc cag gat   162
Leu Leu Gly Ala Leu Leu Gly Thr Ala Trp Ala Arg Arg Ser Gln Asp
    -10   -5   1   5
ctc cac tgt gga gca tgc agg gct ctg gtg gat gaa cta gaa tgg gaa   210
Leu His Cys Gly Ala Cys Arg Ala Leu Val Asp Glu Leu Glu Trp Glu
                10   15   20
att gcc cag gtg gac ccc aag aag acc att cag atg gga tcc ttc cgg   258
Ile Ala Gln Val Asp Pro Lys Lys Thr Ile Gln Met Gly Ser Phe Arg
            25   30   35
atc aat cca gat ggc agc cag tca gtg gtg gag gta act gtt act gtt   306
Ile Asn Pro Asp Gly Ser Gln Ser Val Val Glu Val Thr Val Thr Val
        40   45   50
ccc cca aac aaa gta gct cac tct ggc ttt gga tgaaattcga ctgcttaaaa   359
Pro Pro Asn Lys Val Ala His Ser Gly Phe Gly
    55   60
aggaccttgg tctaatagaa atgaagaaaa cagactcaga aaaaagattt ggctctgtct   419
catttggaag aagctgcagg cttattcccc atgcacttgc ttcctggctg caaaccttaa   479
tactttgttt ctgctgtaga atttgttagc aaacagggag tcctgatcag caccccttctc   539
cacatccaca tgactggttt ttaatgtagc actgtggtat acatgcaaac atccgttcaa   599
aatctgagtc ggagctaaaa ataaaaaatg aaaaaacaaa aaaaaaaaaa   649

<210> 158
<211> 714
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 33..92
<223> Von Heijne matrix
      score 12.3999996185303
      seq ALLLGALLGTAWA/RR
<220>
<221> polyA_site
<222> 703..714
<400> 158

171 300

```
agcagaggtg gagcgacccc attacgctaa ag atg aaa ggc tgg ggt tgg ctg          53
                                      Met Lys Gly Trp Gly Trp Leu
                                      -20                     -15
gcc ctg ctt ctg ggg gcc ctg ctg gga acc gcc tgg gct cgg agg agc         101
Ala Leu Leu Leu Gly Ala Leu Leu Gly Thr Ala Trp Ala Arg Arg Ser
        -10                     -5                      1
cag gat ctc cac tgt gga gca tgc agg gct ctg gtg gat gaa cta gaa         149
Gln Asp Leu His Cys Gly Ala Cys Arg Ala Leu Val Asp Glu Leu Glu
        5                   10                      15
tgg gaa att gcc cag gtg gac ccc aag aag acc att cag atg gga tct         197
Trp Glu Ile Ala Gln Val Asp Pro Lys Lys Thr Ile Gln Met Gly Ser
20                      25                  30                      35
ttc cgg atc aat cca gat ggc agc cag tca gtg gtg gag gtg cct tat         245
Phe Arg Ile Asn Pro Asp Gly Ser Gln Ser Val Val Glu Val Pro Tyr
                    40                  45                  50
gcc cgc tca gag gcc cac ctc aca gag ctg ctg gag gag ata tgt gac         293
Ala Arg Ser Glu Ala His Leu Thr Glu Leu Leu Glu Glu Ile Cys Asp
            55                      60                      65
cgg atg aag gag tat ggg gaa cag att gat cct tcc acc cat cgc aag         341
Arg Met Lys Glu Tyr Gly Glu Gln Ile Asp Pro Ser Thr His Arg Lys
            70                      75                  80
aac tac gta cgt gta gtg ggc cgg aat gga gaa tcc agt gaa ctg gac         389
Asn Tyr Val Arg Val Val Gly Arg Asn Gly Glu Ser Ser Glu Leu Asp
        85                  90                      95
cta caa ggc atc cga atc gac tca gat att agc ggc acc ctc aag ttt.        437
Leu Gln Gly Ile Arg Ile Asp Ser Asp Ile Ser Gly Thr Leu Lys Phe
100                     105                 110                     115
gcg tgt ggg agc att gtg gag gaa tac gag gat gaa ctc att gaa ttc         485
Ala Cys Gly Ser Ile Val Glu Glu Tyr Glu Asp Glu Leu Ile Glu Phe
                    120                 125                     130
ttt tcc cga gag gct gac aat gtt aaa gac aaa ctt tgc agt aag cga         533
Phe Ser Arg Glu Ala Asp Asn Val Lys Asp Lys Leu Cys Ser Lys Arg
            135                 140                     145
aca gat ctt tgt gac cat gcc ctg cac ata tcg cat gat gag cta             578
Thr Asp Leu Cys Asp His Ala Leu His Ile Ser His Asp Glu Leu
            150                 155                     160
tgaaccactg gagcagccca cactggcttg atggatcacc cccaggaggg gaaaatggtg       638
gcaatgcctt ttatatatta tgtttttact gaaattaact gaaaaaatat gaaaccaaaa       698
gtacaaaaaa aaaaaa                                                       714
```

```
<210> 159
<211> 596
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 33..107
<223> Von Heijne matrix
      score 5
      seq MFAASLLAMCAGA/EV
<220>
<221> polyA_signal
<222> 546..551
<220>
<221> polyA_site
<222> 584..596
<400> 159
cacagttcct ctcctcctag agcctgccga cc atg ccc gcg ggc gtg ccc atg        53
```

172

```
                                        Met Pro Ala Gly Val Pro Met
                                        -25                 -20
tcc acc tac ctg aaa atg ttc gca gcc agt ctc ctg gcc atg tgc gca        101
Ser Thr Tyr Leu Lys Met Phe Ala Ala Ser Leu Leu Ala Met Cys Ala
        -15                 -10                 -5
ggg gca gaa gtg gtg cac agg tac tac cga ccg gac ctg aca ata cct        149
Gly Ala Glu Val Val His Arg Tyr Tyr Arg Pro Asp Leu Thr Ile Pro
1                   5                   10
gaa att cca cca aag cgt gga gaa ctc aaa acg gag ctt ttg gga ctg        197
Glu Ile Pro Pro Lys Arg Gly Glu Leu Lys Thr Glu Leu Leu Gly Leu
15                  20                  25                  30
aaa gaa aga aaa cac aaa cct caa gtt tct caa cag gag gaa ctt aaa        245
Lys Glu Arg Lys His Lys Pro Gln Val Ser Gln Gln Glu Glu Leu Lys
                35                  40                  45
taactatgcc aagaattctg tgaataatat aagtcttaaa tatgtatttc ttaatttatt     305
gcatcaaact acttgtcctt aagcacttag tctaatgcta actgcaagag gaggtgctca     365
gtggatgttt agccgatacg ttgaaattta attacggttt gattgatatt tcttgaaaac     425
tgccaaagca catatcatca aaccatttca tgaatatggt ttggaagatg tttagtcttg     485
aatataacgc gaaatagaat atttgtaagt ctactatatg ggttgtcttt atttcatata     545
aattaagaaa ttatttaaaa ctatgaacta gtttcattaa aaaaaaaga a               596


<210> 160
<211> 403
<212> DNA
<213> Homo sapiens
<220>
<221> polyA_signal
<222> 375..380
<220>
<221> polyA_site
<222> 390..403
<400> 160
tgaagagaat ggctgttgca gtcggcgtca gagcagctcc agtgccgggg attcggacgg      60
agagcgcgag gactcggcgg ctgagcgcgc ccgacagcag ctagaggcgc tgctcaacaa     120
gact atg cgc att cgc atg aca gat gga cgg aca ctg gtc ggc tgc ttt      169
     Met Arg Ile Arg Met Thr Asp Gly Arg Thr Leu Val Gly Cys Phe
     1                   5                   10                  15
ctc tgc act gac cgt gac tgc aat gtc atc ctg ggc tcg gcg cag gag      217
Leu Cys Thr Asp Arg Asp Cys Asn Val Ile Leu Gly Ser Ala Gln Glu
                20                  25                  30
ttc ctc aag ccg tcg gat tcc ttc tct gcc ggg gag ccc cgt gtg ctg      265
Phe Leu Lys Pro Ser Asp Ser Phe Ser Ala Gly Glu Pro Arg Val Leu
            35                  40                  45
ggc ctg gcc atg gta ccc gga cac cac atc gtt tcc att gag gtg cag      313
Gly Leu Ala Met Val Pro Gly His His Ile Val Ser Ile Glu Val Gln
        50                  55                  60
agg gag agt ctg acc ggg cct ccg tat ctc tgaccacgat ggcgcttacc        363
Arg Glu Ser Leu Thr Gly Pro Pro Tyr Leu
    65                  70
tttcagactt cattaaactt atgaccaaaa aaaaaaaaa                           403


<210> 161
<211> 727
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 126..575
```

```
ctcagaactg tgctgggaag gatggtaggg cgactggggc tcacctccgc accgttgtag      60
gacccggggt agggttttga gcccgtggga gctgccccac gcggcctcgt cctgccaacg     120
gtcgg atg gcg gag acg aag gac aca gcg cag atg ttg gtg acc ttc aag     170
      Met Ala Glu Thr Lys Asp Thr Ala Gln Met Leu Val Thr Phe Lys
      -150            -145                -140
gat gtg gct gtg acc ttt acc cgg gag gag tgg aga cag ctg gac ctg      218
Asp Val Ala Val Thr Phe Thr Arg Glu Glu Trp Arg Gln Leu Asp Leu
-135            -130                -125                -120
gcc cag agg acc ctg tac cga gag ggc atc ggg ttc ccn aaa cca gag      266
Ala Gln Arg Thr Leu Tyr Arg Glu Gly Ile Gly Phe Pro Lys Pro Glu
                -115                -110                -105
ttg gtc cac ctg cta gag cat ggg cag gag ctg tgg ata gtg aag aga      314
Leu Val His Leu Leu Glu His Gly Gln Glu Leu Trp Ile Val Lys Arg
            -100                -95                 -90
ggc ctc tca cat gct acc tgt gca gag ttt cac tct tgt tgc cca ggc      362
Gly Leu Ser His Ala Thr Cys Ala Glu Phe His Ser Cys Cys Pro Gly
            -85                 -80                 -75
tgg agt gca gtg gnn cgc cat ctc agc tca ctg caa ctt ctg cct ccc      410
Trp Ser Ala Val Xaa Arg His Leu Ser Ser Leu Gln Leu Leu Pro Pro
        -70                 -65                 -60
gag ttc aag gga ttc tcc tgc ctc agc ctc ccg agt agc tgg gat tac      458
Glu Phe Lys Gly Phe Ser Cys Leu Ser Leu Pro Ser Ser Trp Asp Tyr
-55                 -50                 -45                 -40
agg cgc cca cca cca tgc ccg gct ggt ttt ttt gta ttt tta gta gag      506
Arg Arg Pro Pro Pro Cys Pro Ala Gly Phe Phe Val Phe Leu Val Glu
            -35                 -30                 -25
acg ggg ctt cac cat gtt ggc cag gct ggt ctt gaa ctc ttg acc tca      554
Thr Gly Leu His His Val Gly Gln Ala Gly Leu Glu Leu Leu Thr Ser
            -20                 -15                 -10
tgt agt cca ccc gcc tct gcc tcc caa agt gct gcg att aca ggc gtg      602
Cys Ser Pro Pro Ala Ser Ala Ser Gln Ser Ala Ala Ile Thr Gly Val
        -5                  1                   5
agc cac cgt gcc cgg cag aga aaa act gct taaggttgaa aagagaaatt      652
Ser His Arg Ala Arg Gln Arg Lys Thr Ala
10                  15
taagaaattg ctgacggaat aaaaacataa tagaactaca acaccgaagg aaatgaaaga     712
agcaaaaaaa aaaaa                                                       727
```

303

<222> 90..155
<223> Von Heijne matrix
      score 5.90000009536743
      seq IILGCLALFLLLQ/RK
<220>
<221> polyA_signal
<222> 913..918
<220>
<221> polyA_site
<222> 932..944
<400> 162

```
gaatcaggtt ccgtagccca cagaaaagaa gcaagggacg gcaggactgt ttcacacttt        60
tctgcttctg gaaggtgctg acaaaaac atg gaa cta att tcc cca aca gtg         113
                                Met Glu Leu Ile Ser Pro Thr Val
                                         -20                 -15
att ata atc ctg ggt tgc ctt gct ctg ttc tta ctc ctt cag cgg aag        161
Ile Ile Ile Leu Gly Cys Leu Ala Leu Phe Leu Leu Leu Gln Arg Lys
            -10                     -5                      1
aat ttg cgt aga ccc ccg tgc atc aag ggc tgg att cct tgg att gga        209
Asn Leu Arg Arg Pro Pro Cys Ile Lys Gly Trp Ile Pro Trp Ile Gly
         5                      10                  15
gtt gga ttt gag ttt ggg aaa gcc cct cta gaa ttt ata gag aaa gca        257
Val Gly Phe Glu Phe Gly Lys Ala Pro Leu Glu Phe Ile Glu Lys Ala
         20                  25                  30
aga atc aag gta tgt ggt cgt ggc aga cgg ggt ctc cag agg aga caa        305
Arg Ile Lys Val Cys Gly Arg Gly Arg Arg Gly Leu Gln Arg Arg Gln
35                  40                  45                  50
tgc ttt ctt ttt taaactttct ttcattgact cttaagtgca gggctagaac         357
Cys Phe Leu Phe
acggggaaca tacctgcttg cctcaactaa aggatctagt catttctgaa ttcctctact       417
aacaattaac aacaatatcc tgtgcaaaat tttgcgaaag aaatgaaata caattgcagc       477
gtgcatcgac attttggaa gtagagatta acttttcgta tttttacttc atcgaagtta       537
agttccaaat gtgtatgtgt taagtaaatg ttttcagtaa ttgggaaaga taaagtgtaa       597
tccaatttaa gtttgtgaaa atgagtaatt cgtatccaaa ttggagttaa caccaaagta       657
ttgtacaaat tgcttgcaca gttggtccgt acacaataga caggctctgt attttagct        717
gacgttgtta tttgatgatg atgtactcca ttttcactac ggcccgaaga gactagtaat       777
cctccttgta gtagatgttt ttgtcttgaa agtatctttt aaatgtctga gcactttaag       837
gaacagaccc ttattaatgt cttttaagtt ttattcaatt tccagtcaca aatattttat       897
ggtatttgat tgtctaataa atttgtatga tattaaaaaa aaaaaaa                     944
```

<210> 163
<211> 598
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 126..287
<223> Von Heijne matrix
      score 3.90000009536743
      seq LETCGLLVSLVES/IW
<220>
<221> polyA_signal
<222> 561..566
<220>
<221> polyA_site
<222> 587..598
<400> 163

```
ctcagaactg tgctgggaag gatggtaggg cgactggggc tcacctccgc accgttgtag        60
```

175

```
gacccggggt agggtttttga gcccgtggga gctgccccac gcggcctcgt cctgccaacg      120
gtcgg atg gcg gag acg aag gac gca gcg cag atg ttg gtg acc ttc aag      170
      Met Ala Glu Thr Lys Asp Ala Ala Gln Met Leu Val Thr Phe Lys
           -50                 -45                 -40
gat gtg gct gtg acc ttt acc cgg gag gag tgg aga cag ctg gac ctg      218
Asp Val Ala Val Thr Phe Thr Arg Glu Glu Trp Arg Gln Leu Asp Leu
         -35                 -30                 -25
gcc cag agg acc ctg tac cga gag gtg atg ctg gag acc tgt ggg ctt      266
Ala Gln Arg Thr Leu Tyr Arg Glu Val Met Leu Glu Thr Cys Gly Leu
         -20                 -15                 -10
ctg gtt tca cta gtg gaa agc att tgg ctg cat ata aca gaa aac cag      314
Leu Val Ser Leu Val Glu Ser Ile Trp Leu His Ile Thr Glu Asn Gln
         -5                   1                   5
atc aaa ctg gct tca cct gga agg aaa ttc act aac tcg cct gat gag      362
Ile Lys Leu Ala Ser Pro Gly Arg Lys Phe Thr Asn Ser Pro Asp Glu
10                  15                  20                  25
aag cct gag gtg tgg ttg gct cca ggc ctg ttc ggt gcc gca gcc cag      410
Lys Pro Glu Val Trp Leu Ala Pro Gly Leu Phe Gly Ala Ala Ala Gln
                30                  35                  40
tgacgccatc aaggatgtct tggttctctg ttccttcttc ttggttcagg cttctgattg      470
tcctcaggct ggctcctcat agggatgctg ggtgctgcag ccttgactgg ggcagcaggc      530
ccccatgttc aatccatcct cccaccttgg aataaatgct ttcttttcac aatgagaaaa      590
aaaaaaaa                                                                598


<210> 164
<211> 360
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 85..150
<223> Von Heijne matrix
      score 5.90000009536743
      seq IILGCLALFLLLQ/RK
<220>
<221> polyA_site
<222> 349..360
<400> 164
caggttccgt agccacagaa aagaagcaag ggacggcagg actgtttcac acttttctgc       60
ttctggaagg tgctggacaa aaac atg gaa cta att tcc cca aca gtg att      111
                          Met Glu Leu Ile Ser Pro Thr Val Ile
                                   -20                 -15
ata atc ctg ggt tgc ctt gct ctg ttc tta ctc ctt cag cgg aag aat      159
Ile Ile Leu Gly Cys Leu Ala Leu Phe Leu Leu Leu Gln Arg Lys Asn
         -10                 -5                   1
ttg cgt aga ccc ccg tgc atc aag ggc tgg att cct tgg att gga gtt      207
Leu Arg Arg Pro Pro Cys Ile Lys Gly Trp Ile Pro Trp Ile Gly Val
5                   10                  15
gga ttt gag ttt ggg aaa gcc cct cta gaa ttt ata gag aaa gca aga      255
Gly Phe Glu Phe Gly Lys Ala Pro Leu Glu Phe Ile Glu Lys Ala Arg
20                  25                  30                  35
atc aag tat gga cca ata ttt aca gtc ttt gct atg gga aac cga atg      303
Ile Lys Tyr Gly Pro Ile Phe Thr Val Phe Ala Met Gly Asn Arg Met
                40                  45                  50
acc ttt gtt act gaa gaa gaa gga att aat gtg ttt cta aaa tcc      348
Thr Phe Val Thr Glu Glu Glu Gly Ile Asn Val Phe Leu Lys Ser
         55                  60                  65
aaaaaaaaaa aa                                                           360
```

176

<210> 165
<211> 490
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 77..124
<223> Von Heijne matrix
        score 4.80000019073486
        seq SLFIYIFLTCSNT/SP
<220>
<221> polyA_signal
<222> 461..466
<220>
<221> polyA_site
<222> 477..490
<400> 165

```
atgagcttcc agccccaaga gtggaggctg ccacatccca acatagtatc tattgaaaag        60
gaagcagtgt gtatct atg att ata tct ctg ttc atc tat ata ttt ttg aca       112
                   Met Ile Ile Ser Leu Phe Ile Tyr Ile Phe Leu Thr
                   -15                       -10                   -5
tgt agc aac acc tct cca tct tat caa gga act caa ctc ggt ctg ggt         160
Cys Ser Asn Thr Ser Pro Ser Tyr Gln Gly Thr Gln Leu Gly Leu Gly
                 1               5                       10
ctc ccc agt gcc cag tgg tgg cct ttg aca ggt agg agg atg cag tgc         208
Leu Pro Ser Ala Gln Trp Trp Pro Leu Thr Gly Arg Arg Met Gln Cys
        15                  20                  25
tgc agg cta ttt tgt ttt ttg tta caa aac tgt ctt ttc cct ttt ccc         256
Cys Arg Leu Phe Cys Phe Leu Leu Gln Asn Cys Leu Phe Pro Phe Pro
    30                  35                  40
ctc cac ctg att cag cat gat ccc tgt gag ctg gtt ctc aca atc tcc         304
Leu His Leu Ile Gln His Asp Pro Cys Glu Leu Val Leu Thr Ile Ser
45                  50                  55                  60
tgg gac tgg gct gag gca ggg gct tcg ctc tat tct ccc taaccatact          353
Trp Asp Trp Ala Glu Ala Gly Ala Ser Leu Tyr Ser Pro
                65                  70
gtcttccttt ccccccttgcc acttagcagt tatccccccca gctatgcctt ctccctccct     413
cccttgccct ggcatatatt gtgccttatt tatgctgcaa atataacatt aaactatcaa       473
gtgaaaaaaa aaaaaaa                                                       490
```

<210> 166
<211> 488
<212> DNA
<213> Homo sapiens
<220>
<221> polyA_signal
<222> 458..463
<220>
<221> polyA_site
<222> 475..488
<400> 166

```
ccgcttccga aaagagacag acaatgcagc catcata atg aag gtg gac aaa gac         55
                                           Met Lys Val Asp Lys Asp
                                           1               5
cgg cag atg gtg gtg ctg gag gaa gaa ttt cgg aac att tcc cca gag        103
Arg Gln Met Val Val Leu Glu Glu Glu Phe Arg Asn Ile Ser Pro Glu
        10                  15                  20
```

```
gag ctc aaa atg gag ttg ccg gag aga cag ccc agg ttc gtg gtt tac       151
Glu Leu Lys Met Glu Leu Pro Glu Arg Gln Pro Arg Phe Val Val Tyr
            25              30              35
agc tac aag tac gtg cgt gac gat ggc cga gtg tcc tac cct ttg tgt       199
Ser Tyr Lys Tyr Val Arg Asp Asp Gly Arg Val Ser Tyr Pro Leu Cys
        40              45              50
ttc atc ttc tcc agc cct gtg ggc tgc aag ccg gaa caa cag atg atg       247
Phe Ile Phe Ser Ser Pro Val Gly Cys Lys Pro Glu Gln Gln Met Met
55              60              65              70
tat gca ggg agt aaa aac agg ctg gtg cag aca gca gag ctc aca aag       295
Tyr Ala Gly Ser Lys Asn Arg Leu Val Gln Thr Ala Glu Leu Thr Lys
                75              80              85
gtg ttc gaa atc cgc acc act gat gac ctc act gag gcc tgg ctc caa       343
Val Phe Glu Ile Arg Thr Thr Asp Asp Leu Thr Glu Ala Trp Leu Gln
            90              95              100
gaa aag ttg tct ttc ttt cgt tgatctctgg ctggggact gaattcctga          394
Glu Lys Leu Ser Phe Phe Arg
            105
tgtctgagtc ctcaaggtga ctggggactt ggaacccta ggacctgaac aaccaagact      454
ttaaataaat tttaaaatgc aaaaaaaaaa aaaa                                 488


<210> 167
<211> 771
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 48..356
<223> Von Heijne matrix
      score 4.90000009536743
      seq VYAFLGLTAPSGS/KE
<220>
<221> polyA_signal
<222> 742..747
<220>
<221> polyA_site
<222> 760..771
<400> 167
ccacagccct tttcaggacc caaacaaccg cagccgctgt tcccagg atg gtg atc       56
                                                    Met Val Ile
cgt gta tat att gca tct tcc tct ggc tct aca gcg att aag aag aaa      104
Arg Val Tyr Ile Ala Ser Ser Ser Gly Ser Thr Ala Ile Lys Lys Lys
-100            -95             -90             -85
caa caa gat gtg ctt ggt ttc cta gaa gcc aac aaa ata gga ttt gaa      152
Gln Gln Asp Val Leu Gly Phe Leu Glu Ala Asn Lys Ile Gly Phe Glu
            -80             -75             -70
gaa aaa gat att gca gcc aat gaa gag aat cgg aag tgg atg aga gaa      200
Glu Lys Asp Ile Ala Ala Asn Glu Glu Asn Arg Lys Trp Met Arg Glu
            -65             -60             -55
aat gta cct gag aat agt cga cca gcc aca ggt aac ccc ctg cca cct      248
Asn Val Pro Glu Asn Ser Arg Pro Ala Thr Gly Asn Pro Leu Pro Pro
        -50             -45             -40
cag att ttc aat gaa agc cag tat cgc ggg gac tat gat gcc ttc ttt      296
Gln Ile Phe Asn Glu Ser Gln Tyr Arg Gly Asp Tyr Asp Ala Phe Phe
    -35             -30             -25
gaa gcc aga gaa aat aat gca gtg tat gcc ttc tta ggc ttg aca gcc      344
Glu Ala Arg Glu Asn Asn Ala Val Tyr Ala Phe Leu Gly Leu Thr Ala
-20             -15             -10                 -5
```

178 307

```
cca tct ggt tca aag gaa gca gaa gtg caa gca aag cag caa gca          389
Pro Ser Gly Ser Lys Glu Ala Glu Val Gln Ala Lys Gln Gln Ala
              1               5                  10
tgaaccttga gcactgtgct ttaagcatcc tgaaaaatga gtctccattg cttttataaa    449
atagcagaat tagctttgct tcaaaagaaa taggcttaat gttgaaataa tagattagtt    509
gggttttcac atgcaaacac tcaaaatgaa tacaaaatta aaatttgaac attatggtga    569
ttatggtgag gagaatggga tattaacata aaattatatt aataagtaga tatcgtagaa    629
atagtgttgt tacctgccaa gccatcctgt atacaccaat gattttacaa agaaaacacc    689
cttccctcct tctgccatta ctatggcaac ctaagtgtat ctgcagctct acattaaaaa    749
ggagaaagag aaaaaaaaaa aa                                             771


<210> 168
<211> 959
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 69..359
<223> Von Heijne matrix
       score 4
       seq RLPLVVSFIASSS/AN
<220>
<221> polyA_signal
<222> 927..932
<220>
<221> polyA_site
<222> 947..959
<400> 168
cggagagaac caggcagccc agaaacccca ggcgtggaga ttgatcctgc gagagaaggg    60
ggttcatc atg gcg gat gac cta aag cga ttc ttg tat aaa aag tta cca    110
         Met Ala Asp Asp Leu Lys Arg Phe Leu Tyr Lys Lys Leu Pro
              -95             -90                 -85
agt gtt gaa ggg ctc cat gcc att gtt gtg tca gat aga gat gga gta    158
Ser Val Glu Gly Leu His Ala Ile Val Val Ser Asp Arg Asp Gly Val
        -80             -75                 -70
cct gtt att aaa gtg gca aat gac aat gct cca gag cat gct ttg cga    206
Pro Val Ile Lys Val Ala Asn Asp Asn Ala Pro Glu His Ala Leu Arg
        -65             -60                 -55
cct ggt ttc tta tcc act ttt gcc ctt gca aca gac caa gga agc aaa    254
Pro Gly Phe Leu Ser Thr Phe Ala Leu Ala Thr Asp Gln Gly Ser Lys
        -50             -45                 -40
ctt gga ctt tcc aaa aat aaa agt atc atc tgt tac tat aac acc tac    302
Leu Gly Leu Ser Lys Asn Lys Ser Ile Ile Cys Tyr Tyr Asn Thr Tyr
-35                 -30                 -25                 -20
cag gtg gtt caa ttt aat cgt tta cct ttg gtg gtg agt ttc ata gcc    350
Gln Val Val Gln Phe Asn Arg Leu Pro Leu Val Val Ser Phe Ile Ala
        -15                 -10                 -5
agc agc agt gcc aat aca gga cta att gtc agc cta gaa aag gaa ctt    398
Ser Ser Ser Ala Asn Thr Gly Leu Ile Val Ser Leu Glu Lys Glu Leu
        1               5                  10
gct cca ttg ttt gaa gaa ctg aga caa gtt gtg gaa gtt tct            440
Ala Pro Leu Phe Glu Glu Leu Arg Gln Val Val Glu Val Ser
    15                  20                  25
taatctgaca gtggtttcag tgtgtacctt atcttcatta taacaacaca atatcaatcc    500
agcaatcttt agactacaat aatactttta tccatgtgct caagaaaggg cccctttttc    560
caacttatac taaagagcta gcatatagat gtaatttata gatagatcag ttgctatatt    620
ttctggtgta gggtctttct tatttagtga gatctaggga taccacagaa atggttcagt    680
ctatcacagc tcccatggag ttagtctggt caccagatat ggatgagaga ttctattcag    740
```

```
tggatcagaa tcaaactggt acattgatcc acttgagccg ttaagtgctg ccaattgtac      800
aatatgccca ggcttgcaga ataaagccaa ctttttattg tgaataataa taaggacata      860
tttttcttca gattatgttt tatttctttg cattgagtga ggaacataaa atggcttggt      920
aaaagtaata aaatcagtac aatcactaaa aaaaaaaaa                             959


<210> 169
<211> 464
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 33..98
<223> Von Heijne matrix
      score 9.80000019073486
      seq LVVFCLALQLVPG/SP
<220>
<221> polyA_signal
<222> 437..442
<220>
<221> polyA_site
<222> 455..464
<400> 169
gccagaactt actcacccat cccactgaca cc atg aag cct gtg ctg cct ctc        53
                                    Met Lys Pro Val Leu Pro Leu
                                                -20

cag ttc ctg gtg gtg ttc tgc cta gca ctg cag ctg gtg cct ggg agt       101
Gln Phe Leu Val Val Phe Cys Leu Ala Leu Gln Leu Val Pro Gly Ser
-15                     -10             -5                      1

ccc aag cag cgt gtt ctg aag tat atc ttg gaa cct cca ccc tgc ata       149
Pro Lys Gln Arg Val Leu Lys Tyr Ile Leu Glu Pro Pro Pro Cys Ile
            5                   10                  15

tca gca cct gaa aac tgt act cac ctg tgt aca atg cag gaa gat tgc       197
Ser Ala Pro Glu Asn Cys Thr His Leu Cys Thr Met Gln Glu Asp Cys
            20                  25                  30

gag aaa gga ttt cag tgc tgt tcc tcc ttc tgt ggg ata gtc tgt tca       245
Glu Lys Gly Phe Gln Cys Cys Ser Ser Phe Cys Gly Ile Val Cys Ser
      35                  40                  45

tca gaa aca ttt caa aag cgc aac aga atc aaa cac aag ggc tca gaa       293
Ser Glu Thr Phe Gln Lys Arg Asn Arg Ile Lys His Lys Gly Ser Glu
50                  55                  60                  65

gtc atc atg cct gcc aac tgaggcatat ttcctagatc attttgcctc              341
Val Ile Met Pro Ala Asn
                  70

tacgatgttt tttcttggtc cacctttagg aaggtattga gaagcaagaa actggaggcc     401
caatatctaa cctgcaaatc gtttttgagt ttggcaataa aggctaatct accaaaaaaa     461
aaa                                                                    464


<210> 170
<211> 799
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 110..235
<223> Von Heijne matrix
      score 5.19999980926514
      seq LLFDLVCHEFCQS/DD
<220>
```

<221> polyA_signal
<222> 764..769
<220>
<221> polyA_site
<222> 787..799
<400> 170
ccaaccccag gaagagtctg aagagcagcc agtgtttcgg cttgtgccct gtatacttga   60
agctgccaaa caagtacgtt ctgaaaatcc agaatggctt gatgtttac atg cac att   118
                                                        Met His Ile
                                                             -40
tta caa ctg ctt act aca gtg gat gat gga att caa gca att gta cat   166
Leu Gln Leu Leu Thr Thr Val Asp Asp Gly Ile Gln Ala Ile Val His
            -35   -30   -25
tgt cct gac act gga aaa gac att tgg aat cta ctt ttt gac ctg gtc   214
Cys Pro Asp Thr Gly Lys Asp Ile Trp Asn Leu Leu Phe Asp Leu Val
        -20   -15   -10
tgc cat gaa ttc tgc cag tct gat gat cca ccc atc att ctt caa gaa   262
Cys His Glu Phe Cys Gln Ser Asp Asp Pro Pro Ile Ile Leu Gln Glu
        -5   1   5
cag aaa aca gtg cta gcc tct gtt ttt tca gtg ttg tct gcc atc tat   310
Gln Lys Thr Val Leu Ala Ser Val Phe Ser Val Leu Ser Ala Ile Tyr
10   15   20   25
gcc tca cag act gag caa gag tat cta aag ata gaa aaa gta gat ctt   358
Ala Ser Gln Thr Glu Gln Glu Tyr Leu Lys Ile Glu Lys Val Asp Leu
                30   35   40
cct cta att gac agc ctc att cgg gtc tta caa aat atg gaa cag tgt   406
Pro Leu Ile Asp Ser Leu Ile Arg Val Leu Gln Asn Met Glu Gln Cys
                45   50   55
cag aaa aaa cca gag aac tcg gca gag tct aac aca gag gaa act aaa   454
Gln Lys Lys Pro Glu Asn Ser Ala Glu Ser Asn Thr Glu Glu Thr Lys
            60   65   70
agg act gat tta acc caa gat gat ttc cac ttg aaa atc tta aag gat   502
Arg Thr Asp Leu Thr Gln Asp Asp Phe His Leu Lys Ile Leu Lys Asp
        75   80   85
att tta tgt gaa ttt ctt tct aat att ttt cag gca tta aca aag gag   550
Ile Leu Cys Glu Phe Leu Ser Asn Ile Phe Gln Ala Leu Thr Lys Glu
90   95   100   105
acg gtg gct cag gga gta aag gaa ggc cag ttg agc aaa cag aag tgt   598
Thr Val Ala Gln Gly Val Lys Glu Gly Gln Leu Ser Lys Gln Lys Cys
                110   115   120
tcc tct gca ttt caa aac ctt ctt cct ttc tat agc cct gtg gtg gaa   646
Ser Ser Ala Phe Gln Asn Leu Leu Pro Phe Tyr Ser Pro Val Val Glu
                125   130   135
gat ttt att aaa atc cta cgt gaa gtt gat aag gcg ctt gct gat gac   694
Asp Phe Ile Lys Ile Leu Arg Glu Val Asp Lys Ala Leu Ala Asp Asp
            140   145   150
ttg gaa aaa aac ttc cca agt ttg aag gtt cag act taaaacctga   740
Leu Glu Lys Asn Phe Pro Ser Leu Lys Val Gln Thr
            155   160   165
attggaatta cttctgtaca agaaataaac tttattttc tcactgaaaa aaaaaaaaa   799

<210> 171
<211> 320
<212> DNA
<213> Homo sapiens
<220>
<221> polyA_site
<222> 308..320

181

```
<400> 171
tcatcatcca gagcagccag tgtccgggag gcagaag atg ccc cac tcc aag cct         55
                                           Met Pro His Ser Lys Pro
                                           1                   5
ctg gac tgg ggg ctc tct tca gtg gct gaa tgt cca gca gag cta ttt         103
Leu Asp Trp Gly Leu Ser Ser Val Ala Glu Cys Pro Ala Glu Leu Phe
            10                  15                  20
cct tcc aca ggg ggc ctt gca ggg aag ggt cca gga ctt gac atc tta         151
Pro Ser Thr Gly Gly Leu Ala Gly Lys Gly Pro Gly Leu Asp Ile Leu
        25                  30                  35
aga tgc gtc ttg tcc cct tgg gcc agt cat ttc ccc tct ctg agc ctc         199
Arg Cys Val Leu Ser Pro Trp Ala Ser His Phe Pro Ser Leu Ser Leu
    40                  45                  50
ggt gtc ttc aac ctg tgaaatggga tcataatcac tgccttacct ccctcacggt         254
Gly Val Phe Asn Leu
55
tgttgtgagg actgagtgtg tggaagtttt tcataaactt tggatgctag tgtaaaaaaa        314
aaaaaa                                                                   320


<210> 172
<211> 331
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 129..209
<223> Von Heijne matrix
      score 4.90000009536743
      seq CLLSYIALGAIHA/KI
<220>
<221> polyA_site
<222> 318..331
<400> 172
atggaaacca gatggggcaa cggggtggtt ctagtgcaga ctgtagctgc agctcctctc         60
cacctctagc ctgctcattt ccagctcaga aattctacta atggcgtttt ttcttcctga        120
aaaaggaa atg aac agg gtc cct gct gat tct cca aat atg tgt cta atc        170
        Met Asn Arg Val Pro Ala Asp Ser Pro Asn Met Cys Leu Ile
            -25                 -20                 -15
tgt tta ctg agt tac ata gca ctt gga gcc atc cat gca aaa atc tgt        218
Cys Leu Leu Ser Tyr Ile Ala Leu Gly Ala Ile His Ala Lys Ile Cys
        -10                 -5                  1
agg aga gca ttc cag gaa gag gga aga gca aat gca aag acg ggc gtg        266
Arg Arg Ala Phe Gln Glu Glu Gly Arg Ala Asn Ala Lys Thr Gly Val
    5                   10                  15
aga gct tgg tgc ata cag cca tgg gcc aaa taaagtttcc ttggaatagc        316
Arg Ala Trp Cys Ile Gln Pro Trp Ala Lys
20                  25
caaaaaaaaa aaaaa                                                         331


<210> 173
<211> 1075
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 78..359
<223> Von Heijne matrix
      score 4.19999980926514
```

311

```
gtggtaggga gcagccagga gcggttttct gggaactgtg ggatgtgccc ttgggggccc      60
gagaaaacag aaggaag atg ctc cag acc agt aac tac agc ctg gtg ctc         110
                    Met Leu Gln Thr Ser Asn Tyr Ser Leu Val Leu
                            -90                         -85
tct ctg cag ttc ctg ctg ctg tcc tat gac ctc ttt gtc aat tcc ttc       158
Ser Leu Gln Phe Leu Leu Leu Ser Tyr Asp Leu Phe Val Asn Ser Phe
            -80                     -75                     -70
tca gaa ctg ctc caa aag act cct gtc atc cag ctt gtg ctc ttc atc       206
Ser Glu Leu Leu Gln Lys Thr Pro Val Ile Gln Leu Val Leu Phe Ile
        -65                     -60                     -55
atc cag gat att gca gtc ctc ttc aac atc atc atc att ttc ctc atg       254
Ile Gln Asp Ile Ala Val Leu Phe Asn Ile Ile Ile Ile Phe Leu Met
    -50                     -45                     -40
ttc ttc aac acc ttc gtc ttc cag gct ggc ctg gtc aac ctc cta ttc       302
Phe Phe Asn Thr Phe Val Phe Gln Ala Gly Leu Val Asn Leu Leu Phe
-35                     -30                     -25                 -20
cat aag ttc aaa ggg acc atc atc ctg aca gct gtg tac ttt gcc ctc       350
His Lys Phe Lys Gly Thr Ile Ile Leu Thr Ala Val Tyr Phe Ala Leu
            -15                     -10                     -5
agc atc tcc ctt cat gtc tgg gtc atg aac tta cgc tgg aaa aac tcc       398
Ser Ile Ser Leu His Val Trp Val Met Asn Leu Arg Trp Lys Asn Ser
            1                   5                       10
aac agc ttc ata tgg aca gat gga ctt caa atg ctg ttt gta ttc cag       446
Asn Ser Phe Ile Trp Thr Asp Gly Leu Gln Met Leu Phe Val Phe Gln
    15                      20                      25
aga cta gca gca gtg ttg tac tgc tac ttc tat aaa cgg aca gcc gta       494
Arg Leu Ala Ala Val Leu Tyr Cys Tyr Phe Tyr Lys Arg Thr Ala Val
30                      35                      40                  45
aga cta ggc gat cct cac ttc tac cag gac tct ttg tgg ctg cgc aag       542
Arg Leu Gly Asp Pro His Phe Tyr Gln Asp Ser Leu Trp Leu Arg Lys
            50                      55                      60
gag ttc atg caa gtt cga agg tgacctcttg tcacactgat ggatactttt          593
Glu Phe Met Gln Val Arg Arg
            65
ccttcctgat agaagccaca tttgctgctt tgcagggaga gttggcccta tgcatgggca      653
aacagctgga ctttccaagg aaggttcaga ctagctgtgt tcagcattca agaaggaaga      713
tccccctct tgcacaatta gagtgtcccc atcggtctcc agtgcggcat cccttccttg       773
ccttctacct ctgttccacc cccttccttc ctctcctctc tgtaccattc attctccctg      833
accggccttt cttgccgagg gttctgtggc tcttacctt gtgaagcttt tcctttagcc       893
tgggacagaa ggacctcccg gcccccaaag gatctcccag tgaccaaagg atgcgaagag      953
tgatagttac gtgctcctga ctgatcacac cgcagacatt tagattttta tacccaaggc     1013
actttaaaaa aatgttttat aaatagagaa taaattgaat tcttgttcca aaaaaaaaaa     1073
aa                                                                   1075
```

```
cactgggtca aggagtaagc agaggataaa caactggaag gagagcaagc acaaagtcat      60
c atg gct tca gcg tct gct cgt gga aac caa gat aaa gat gcc cat ttt     109
  Met Ala Ser Ala Ser Ala Arg Gly Asn Gln Asp Lys Asp Ala His Phe
            -65                 -60                 -55
cca cca cca agc aag cag agc ctg ttg ttt tgt cca aaa tca aaa ctg       157
Pro Pro Pro Ser Lys Gln Ser Leu Leu Phe Cys Pro Lys Ser Lys Leu
        -50                 -45                 -40
cac atc cac aga gca gag atc tca aag att atg cga gaa tgt cag gaa       205
His Ile His Arg Ala Glu Ile Ser Lys Ile Met Arg Glu Cys Gln Glu
    -35                 -30                 -25
gaa agt ttc tgg aag aga gct ctg cct ttt tct ctt gta agc atg ctt       253
Glu Ser Phe Trp Lys Arg Ala Leu Pro Phe Ser Leu Val Ser Met Leu
-20                 -15                 -10                 -5
gtc acc cag gga cta gtc tac caa ggt tat ttg gca gct aat tct aga       301
Val Thr Gln Gly Leu Val Tyr Gln Gly Tyr Leu Ala Ala Asn Ser Arg
                1                   5                   10
ttt gga tca ttg ccc aaa gtt gca ctt gct ggt ctc ttg gga ttt ggc       349
Phe Gly Ser Leu Pro Lys Val Ala Leu Ala Gly Leu Leu Gly Phe Gly
        15                  20                  25
ctt gga aag gta tca tac ata gga gta tgc cag agt aaa ttc cat ttt       397
Leu Gly Lys Val Ser Tyr Ile Gly Val Cys Gln Ser Lys Phe His Phe
        30                  35                  40
ttt gaa gat cag ctc cgt ggg gct ggt ttt ggt cca cag cat aac agg       445
Phe Glu Asp Gln Leu Arg Gly Ala Gly Phe Gly Pro Gln His Asn Arg
45                  50                  55                  60
cac tgc ctc ctt acc tgt gag gaa tgc aaa ata aag cat gga tta agt       493
His Cys Leu Leu Thr Cys Glu Glu Cys Lys Ile Lys His Gly Leu Ser
                65                  70                  75
gag aag gga gac tct cag cct tca gct tcc taaattctgt gtctgtgact         543
Glu Lys Gly Asp Ser Gln Pro Ser Ala Ser
                80                  85
ttcgaagttt tttaaacctc tgaatttgta cacatttaaa atttcaagtg tactttaaaa     603
taaaatactt ctaatgtaaa aaaaaaaaa                                        632
```

```
gtattgggaa agtgatttgt gaa atg aaa gta gaa gaa gag cat acc aat gca      53
                      Met Lys Val Glu Glu Glu His Thr Asn Ala
```

```
                        1                   5                              10
ata ggc act ctc cac ggc ggt ttg aca gcc acg tta gta gat aac ata        101
Ile Gly Thr Leu His Gly Gly Leu Thr Ala Thr Leu Val Asp Asn Ile
                15                  20                  25
tca aca atg gct ctg cta tgc acg gaa agg gga gca ccc gga gtc agt        149
Ser Thr Met Ala Leu Leu Cys Thr Glu Arg Gly Ala Pro Gly Val Ser
            30                  35                  40
gtc gat atg aac ata acg tac atg tca cct gca aaa tta gga gag gat        197
Val Asp Met Asn Ile Thr Tyr Met Ser Pro Ala Lys Leu Gly Glu Asp
            45                  50                  55
ata gtg att aca gca cat gtt ctg aag caa gga aaa aca ctt gca ttt        245
Ile Val Ile Thr Ala His Val Leu Lys Gln Gly Lys Thr Leu Ala Phe
    60                  65                  70
acc tct gtg ggt ctg acc aac aag gcc aca gga aaa tta ata gca caa        293
Thr Ser Val Gly Leu Thr Asn Lys Ala Thr Gly Lys Leu Ile Ala Gln
75                  80                  85                  90
gga aga cac aca aaa cac ctg gga aac tgagagaaca gcagaatgac            340
Gly Arg His Thr Lys His Leu Gly Asn
                95

ctaaagaaac ccaacaatga atatcaagta tagatttgac tcaaacaatt gtaatttttg    400
aaataaacta gcaaaccaa aaaaaaaaaa                                      430
```

<210> 176
<211> 185
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 42..113
<223> Von Heijne matrix
      score 3.70000004768372
      seq ILFNLLIFLCGFT/NY
<220>
<221> polyA_site
<222> 172..185
<400> 176

```
ctttcagaac tcactgccaa gagccctgaa caggagccac c atg cag tgc ttc agc    56
                                              Met Gln Cys Phe Ser
                                                              -20
ttc att aag acc atg atg atc ctc ttc aat ttg ctc atc ttt ctg tgt     104
Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu Leu Ile Phe Leu Cys
                -15                 -10                 -5
ggc ttc acc aac tat acg gat ttt gag gac tca ccc tac ttc aaa atg     152
Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp Ser Pro Tyr Phe Lys Met
            1                   5                   10
cat aaa cct gtt aca atg taaaaaaaaa aaaaa                             185
His Lys Pro Val Thr Met
    15
```

<210> 177
<211> 585
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 108..170
<223> Von Heijne matrix
      score 5.5

```
          seq SFLPSALVIWTSA/AF
<220>
<221> polyA_signal
<222> 550..555
<220>
<221> polyA_site
<222> 574..585
<400> 177
cacgttcctg ttgagtacac gttcctgttg atttacaaaa ggtgcaggta tgagcaggtc        60
tgaagactaa cattttgtga agttgtaaaa cagaaaacct gttagaa atg tgg tgg        116
                                                        Met Trp Trp
                                                        -20
ttt cag caa ggc ctc agt ttc ctt cct tca gcc ctt gta att tgg aca        164
Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr
         -15                 -10                 -5
tct gct gct ttc ata ttt tca tac att act gca gta aca ctc cac cat        212
Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His
    1               5                   10
ata gac ccg gct tta cct tat atc agt gac act ggt aca gta gct cca        260
Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro
15                  20                  25                  30
gaa aaa tgc tta ttt ggg gca atg cta aat att gcg gca gtc tta tgt        308
Glu Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys
            35                  40                  45
caa aaa tagaaatcag gaagataatt caacttaaag aagttcattt catgaccaaa        364
Gln Lys
ctcttcagaa acatgtcttt acaagcatat ctcttgtatt gctttctaca ctgttgaatt        424
gtctggcaat atttctgcag tggaaaattt gatttagcta gttcttgact gataaatatg        484
gtaaggtggg cttttccccc tgtgtaattg gctactatgt cttactgagc caagttgtaa        544
tttgaaataa aatgatatga gagtgacaca aaaaaaaaaa a                            585


<210> 178
<211> 613
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 118..171
<223> Von Heijne matrix
        score 5.90000009536743
        seq ALALLWSLPASDL/GR
<220>
<221> polyA_signal
<222> 583..588
<220>
<221> polyA_site
<222> 602..613
<400> 178
ggggtgggtg gactagaagc atttgggagt agtggccagg ggccctggac gctagccacg        60
gagctgccgc acagagcctg gtgtccacaa gcttccaggt tggggttgga gcctggg        117
atg agc ccc ggc agc gcc ttg gcc ctt ctg tgg tcc ctg cca gcc tct        165
Met Ser Pro Gly Ser Ala Leu Ala Leu Leu Trp Ser Leu Pro Ala Ser
            -15                 -10                 -5
gac ctg ggc cgg tca gtc att gct gga ctc tgg cca cac act ggc gtt        213
Asp Leu Gly Arg Ser Val Ile Ala Gly Leu Trp Pro His Thr Gly Val
    1               5                   10
ctc atc cac ttg gaa aca agc cag tct ttt ctg caa ggt cag ttg acc        261
Leu Ile His Leu Glu Thr Ser Gln Ser Phe Leu Gln Gly Gln Leu Thr
```

```
          15                    20                    25                    30
aag agc ata ttt ccc ctc tgt tgt aca tcg ttg ttt tgt gtt tgt gtt          309
Lys Ser Ile Phe Pro Leu Cys Cys Thr Ser Leu Phe Cys Val Cys Val
                         35                    40                    45
gta aca gtg ggt gga ggg agg gtg ggg tct aca ttt gtt gca                  351
Val Thr Val Gly Gly Gly Arg Val Gly Ser Thr Phe Val Ala
                    50                    55                    60
tgagtcgatg ggtcagaact ttagtatacg catgcgtcct ctgagtgaca gggcattttg        411
tcgaaaataa gcaccttggt aactaaaccc ctctaatagc tataaaggct ttagttctgt        471
attgattaag ttactgtaaa agcttgggtt tatttttgta ggacttaatg ctaagaatt        531
agaacatagc aagggggctc ctctgttgga gtaatgtaaa ttgtaattat aaataaacat        591
gcaaaccttt aaaaaaaaaa aa                                                  613

<210> 179
<211> 427
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 128..268
<223> Von Heijne matrix
      score 5.5
      seq SALLFFARPCVFC/FK
<220>
<221> polyA_signal
<222> 410..415
<220>
<221> polyA_site
<222> 424..427
<400> 179
agcttggatt tacactgggc aacgtggttg gaatgtatct ggctcagaac tatgatatac         60
caaacctggc taaaaaactt gaagaaatta aaaaggactt ggatgccaag aagaaacccc        120
ctagtgc atg aga ctg cct cca gca ctg cct tca gga tat act gat tct          169
        Met Arg Leu Pro Pro Ala Leu Pro Ser Gly Tyr Thr Asp Ser
            -45                   -40                   -35
act gct ctt gag ggc ctc gtt tac tat ctg aac caa aag ctt ttg ttt          217
Thr Ala Leu Glu Gly Leu Val Tyr Tyr Leu Asn Gln Lys Leu Leu Phe
            -30                   -25                   -20
tcg tct cca gcc tca gca ctt ctc ttc ttt gct aga ccc tgt gtt ttt          265
Ser Ser Pro Ala Ser Ala Leu Leu Phe Phe Ala Arg Pro Cys Val Phe
            -15                   -10                   -5
tgc ttt aaa gca agc aaa atg ggg ccc caa ttt gag aac tac cca aca          313
Cys Phe Lys Ala Ser Lys Met Gly Pro Gln Phe Glu Asn Tyr Pro Thr
    1                   5                     10                    15
ttt cca aca tac tca cct ctt ccc ata atc cct ttc caa ctg cat ggg          361
Phe Pro Thr Tyr Ser Pro Leu Pro Ile Ile Pro Phe Gln Leu His Gly
                    20                    25                    30
agg ttc taagactgga attatggtgc tagattagta aacatgactt ttaatgaaaa           417
Arg Phe
aaaaacaaaa                                                                427

<210> 180
<211> 905
<212> DNA
<213> Homo sapiens
<220>
<221> sig_peptide
<222> 149..457
```

316

```
gctgcctgtt cttcacactt agctccaaac ccatgaaaaa ttgccaagta taaaagcttc      60
tcaagaatga gatggattct agggtgtctt cacctgagaa gcaagataaa gagaatttcg     120
tgggtgtcaa caataaacgg cttggtgt atg tgg ctg gat cct gtt ttc cct       172
                                 Met Trp Leu Asp Pro Val Phe Pro
                                                         -100
ctc ttt cct gtt ggt gat cat tac ctt ccc cat ctc cat atg gat gtg     220
Leu Phe Pro Val Gly Asp His Tyr Leu Pro His Leu His Met Asp Val
-95             -90             -85                     -80
ctt gaa ggt ttg atc ctg gtc ctg cca tgc ata gat gtg ttt gtc aaa     268
Leu Glu Gly Leu Ile Leu Val Leu Pro Cys Ile Asp Val Phe Val Lys
                -75                 -70                 -65
gtt gac ctc cga aca gtt act tgc aac att cct cca caa gag atc ctc     316
Val Asp Leu Arg Thr Val Thr Cys Asn Ile Pro Pro Gln Glu Ile Leu
            -60                 -55                 -50
acc aga gac tcc gta act act cag gta gat gga gtt gtc tat tac aga     364
Thr Arg Asp Ser Val Thr Thr Gln Val Asp Gly Val Val Tyr Tyr Arg
            -45                 -40                 -35
atc tat agt gct gtc tca gca gtg gct aat gtc aac gat gtc cat caa     412
Ile Tyr Ser Ala Val Ser Ala Val Ala Asn Val Asn Asp Val His Gln
    -30                 -25                 -20
gca aca ttt ctg ctg gct caa acc act ctg aga aat gtc tta ggg aca     460
Ala Thr Phe Leu Leu Ala Gln Thr Thr Leu Arg Asn Val Leu Gly Thr
-15                 -10                 -5                  1
cag acc ttg tcc cag atc tta gct gga cga gaa gag atc gcc cat agc     508
Gln Thr Leu Ser Gln Ile Leu Ala Gly Arg Glu Glu Ile Ala His Ser
            5                   10                  15
atc cag act tta ctt gat gat gcc acc gaa ctg tgg ggg atc cgg gtg     556
Ile Gln Thr Leu Leu Asp Asp Ala Thr Glu Leu Trp Gly Ile Arg Val
            20                  25                  30
gcc cga gtg gaa atc aaa gat gtt cgg att ccc gtg cag ttg cag aga     604
Ala Arg Val Glu Ile Lys Asp Val Arg Ile Pro Val Gln Leu Gln Arg
    35                  40                  45
tcc atg gca gcc gag gct gag gcc acc cgg gaa gcg aga gcc aag gtc     652
Ser Met Ala Ala Glu Ala Glu Ala Thr Arg Glu Ala Arg Ala Lys Val
50                  55                  60                  65
ctt gca gct gaa gga gaa atg aat gct tcc aaa tcc ctg aag tca gcc     700
Leu Ala Ala Glu Gly Glu Met Asn Ala Ser Lys Ser Leu Lys Ser Ala
                70                  75                  80
tcc atg gtg ctg gct gag tct ccc ata gct ctc cag ctg cgc tac ctg     748
Ser Met Val Leu Ala Glu Ser Pro Ile Ala Leu Gln Leu Arg Tyr Leu
                85                  90                  95
cag acc ttg agc acg gta gcc acc gag aag aat tct acg att gtg ttt     796
Gln Thr Leu Ser Thr Val Ala Thr Glu Lys Asn Ser Thr Ile Val Phe
                100                 105                 110
cct ctg ccc atg aat ata cta gag ggc att ggt ggc gtc agc tat gat     844
Pro Leu Pro Met Asn Ile Leu Glu Gly Ile Gly Gly Val Ser Tyr Asp
    115                 120                 125
aac cac aag aag ctt cca aat aaa gcc tgaggtcctc ttgcggtagt            891
Asn His Lys Lys Leu Pro Asn Lys Ala
130                 135
caaaaaaaaa aaaa                                                       905
```

188

<210> 181
<211> 307
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -13..-1
<400> 181

```
Met Leu Ala Val Ser Leu Thr Val Pro Leu Leu Gly Ala Met Met Leu
            -10                  -5                   1
Leu Glu Ser Pro Ile Asp Pro Gln Pro Leu Ser Phe Lys Glu Pro Pro
      5                  10                  15
Leu Leu Leu Gly Val Leu His Pro Asn Thr Lys Leu Arg Gln Ala Glu
20                  25                  30                  35
Arg Leu Phe Glu Asn Gln Leu Val Gly Pro Glu Ser Ile Ala His Ile
                  40                  45                  50
Gly Asp Val Met Phe Thr Gly Thr Ala Asp Gly Arg Val Val Lys Leu
              55                  60                  65
Glu Asn Gly Glu Ile Glu Thr Ile Ala Arg Phe Gly Ser Gly Pro Cys
          70                  75                  80
Lys Thr Arg Gly Asp Glu Pro Val Cys Gly Arg Pro Leu Gly Ile Arg
      85                  90                  95
Ala Gly Pro Asn Gly Thr Leu Phe Val Ala Asp Ala Tyr Lys Gly Leu
100                 105                 110                 115
Phe Glu Val Asn Pro Trp Lys Arg Glu Val Lys Leu Leu Leu Ser Ser
                  120                 125                 130
Glu Thr Pro Ile Glu Gly Lys Asn Met Ser Phe Val Asn Asp Leu Thr
              135                 140                 145
Val Thr Gln Asp Gly Arg Lys Ile Tyr Phe Thr Asp Ser Ser Ser Lys
          150                 155                 160
Trp Gln Arg Arg Asp Tyr Leu Leu Leu Val Met Glu Gly Thr Asp Asp
      165                 170                 175
Gly Arg Leu Leu Glu Tyr Asp Thr Val Thr Arg Glu Val Lys Val Leu
180                 185                 190                 195
Leu Asp Gln Leu Arg Phe Pro Asn Gly Val Gln Leu Ser Pro Ala Glu
                  200                 205                 210
Asp Phe Val Leu Val Ala Glu Thr Thr Met Ala Arg Ile Arg Arg Val
              215                 220                 225
Tyr Val Ser Gly Leu Met Lys Gly Gly Ala Asp Leu Phe Val Glu Asn
          230                 235                 240
Met Pro Gly Phe Pro Asp Asn Ile Arg Pro Ser Ser Ser Gly Gly Tyr
      245                 250                 255
Trp Val Gly Met Ser Thr Ile Arg Pro Asn Pro Gly Phe Ser Met Leu
260                 265                 270                 275
Asp Phe Leu Ser Glu Arg Pro Trp Ile Lys Arg Met Ile Phe Lys Val
                  280                 285                 290
Lys Lys Lys
```

<210> 182
<211> 59
<212> PRT
<213> Homo sapiens
<400> 182

```
Met Met Tyr Val Ser Ile Glu Met Ser Gly Pro Thr Ile Ser His Leu
1                   5                   10                  15
Phe Asp Tyr Val Val Cys Tyr Ile Tyr Gly Leu Lys Ser Phe Ser Leu
          20                  25                  30
```

```
Lys Gln Leu Lys Lys Lys Ser Trp Ser Lys Tyr Leu Phe Glu Ser Cys
        35              40              45
Cys Tyr Arg Ser Leu Tyr Val Cys Val Phe Ile
    50              55


<210> 183
<211> 97
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -28..-1
<400> 183
Met Ser Pro Ala Phe Arg Ala Met Asp Val Glu Pro Arg Ala Lys Gly
        -25             -20             -15
Val Leu Leu Glu Pro Phe Val His Gln Val Gly Gly His Ser Cys Val
        -10             -5              1
Leu Arg Phe Asn Glu Thr Thr Leu Cys Lys Pro Leu Val Pro Arg Glu
5               10              15              20
His Gln Phe Tyr Glu Thr Leu Pro Ala Glu Met Arg Lys Phe Ser Pro
            25              30              35
Gln Tyr Lys Gly Gln Ser Gln Arg Pro Leu Val Ser Trp Pro Ser Leu
        40              45              50
Pro His Phe Phe Pro Trp Ser Phe Pro Leu Trp Pro Gln Gly Ser Val
        55              60              65
Ala


<210> 184
<211> 52
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -32..-1
<400> 184
Met Leu Gly Thr Thr Gly Leu Gly Thr Gln Gly Pro Ser Gln Gln Ala
        -30             -25             -20
Leu Gly Phe Phe Ser Phe Met Leu Leu Gly Met Gly Gly Cys Leu Pro
        -15             -10             -5
Gly Phe Leu Leu Gln Pro Pro Asn Arg Ser Pro Thr Leu Pro Ala Ser
1               5                       10              15
Thr Phe Ala His
            20


<210> 185
<211> 124
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -97..-1
<400> 185
Met Ala Asp Asp Leu Lys Arg Phe Leu Tyr Lys Lys Leu Pro Ser Val
        -95             -90             -85
Glu Gly Leu His Ala Ile Val Val Ser Asp Arg Asp Gly Val Pro Val
        -80             -75             -70
Val Lys Val Ala Asn Asp Asn Ala Pro Glu His Ala Leu Arg Pro Gly
-65             -60             -55             -50
```

```
Phe Leu Ser Thr Phe Ala Leu Ala Thr Asp Gln Gly Ser Lys Leu Gly
             -45                 -40                     -35
Leu Ser Lys Asn Lys Ser Ile Ile Cys Tyr Tyr Asn Thr Tyr Gln Val
         -30                 -25                 -20
Val Gln Phe Asn Arg Leu Pro Leu Val Val Ser Phe Ile Ala Ser Ser
         -15                 -10                 -5
Ser Ala Asn Thr Gly Leu Ile Val Ser Leu Glu Lys Glu Leu Ala Pro
     1               5               10                      15
Leu Phe Glu Glu Leu Arg Gln Val Val Glu Val Ser
                 20                 25
```

<210> 186
<211> 230
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -24..-1
<400> 186

```
Met Ala Ser Leu Gly Leu Gln Leu Val Gly Tyr Ile Leu Gly Leu Leu
             -20                 -15                 -10
Gly Leu Leu Gly Thr Leu Val Ala Met Leu Leu Pro Ser Trp Lys Thr
         -5                  1               5
Ser Ser Tyr Val Gly Ala Ser Ile Val Thr Ala Val Gly Phe Ser Lys
     10              15                  20
Gly Leu Trp Met Glu Cys Ala Thr His Ser Thr Gly Ile Thr Gln Cys
25                  30                  35                  40
Asp Ile Tyr Ser Thr Leu Leu Gly Leu Pro Ala Asp Ile Gln Ala Ala
             45                  50                  55
Gln Ala Met Met Val Thr Ser Ser Ala Ile Ser Ser Leu Ala Cys Ile
             60                  65                  70
Ile Ser Val Val Gly Met Arg Cys Thr Val Phe Cys Gln Glu Ser Arg
         75                  80                  85
Ala Lys Asp Arg Val Ala Val Ala Gly Gly Val Phe Phe Ile Leu Gly
         90                  95                  100
Gly Leu Leu Gly Phe Ile Pro Val Ala Trp Asn Leu His Gly Ile Leu
105                 110                 115                 120
Arg Asp Phe Tyr Ser Pro Leu Val Pro Asp Ser Met Lys Phe Glu Ile
             125                 130                 135
Gly Glu Ala Leu Tyr Leu Gly Ile Ile Ser Ser Leu Phe Ser Leu Ile
             140                 145                 150
Ala Gly Ile Ile Leu Cys Phe Ser Cys Ser Ser Gln Arg Asn Arg Ser
         155                 160                 165
Asn Tyr Tyr Asp Ala Tyr Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser
     170                 175                 180
Pro Arg Pro Gly Gln Pro Pro Lys Val Lys Ser Glu Phe Asn Ser Tyr
185                 190                 195                 200
Ser Leu Thr Gly Tyr Val
             205
```

<210> 187
<211> 72
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -32..-1
<400> 187

```
Met Phe Ala Leu Ala Val Met Arg Ala Phe Arg Lys Asn Lys Thr Leu
        -30                 -25                 -20
Gly Tyr Gly Val Pro Met Leu Leu Ile Ala Gly Gly Ser Phe Gly
        -15             -10                 -5
Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp Ala Val Lys Ser Lys Met
1                   5                   10                  15
Asp Pro Glu Leu Glu Lys Lys Pro Lys Glu Asn Lys Ile Ser Leu Glu
            20                  25                  30
Ser Glu Tyr Glu Gly Ser Ile Cys
            35              40
```

<210> 188
<211> 88
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -33..-1
<400> 188

```
Met Ser Gln Thr Ala Trp Leu Ser Leu Leu Ser Ser Ser Pro Phe Gly
        -30                 -25                 -20
Pro Phe Ser Ala Leu Thr Phe Leu Phe Leu His Leu Pro Pro Ser Thr
        -15             -10                 -5
Ser Leu Phe Ile Asn Leu Ala Arg Gly Gln Ile Lys Gly Pro Leu Gly
    1               5                   10                  15
Leu Ile Leu Leu Leu Ser Phe Cys Gly Gly Tyr Thr Lys Cys Asp Phe
            20                  25                  30
Ala Leu Ser Tyr Leu Glu Ile Pro Asn Arg Ile Glu Phe Ser Ile Met
            35              40                  45
Asp Pro Lys Arg Lys Thr Lys Cys
            50              55
```

<210> 189
<211> 106
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -32..-1
<400> 189

```
Met Phe Ala Pro Ala Val Thr Arg Ala Phe Arg Lys Asn Lys Thr Leu
        -30                 -25                 -20
Gly Tyr Gly Val Pro Met Leu Leu Leu Ile Val Gly Gly Ser Phe Gly
        -15             -10                 -5
Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp Ala Val Lys Ser Lys Met
1                   5                   10                  15
Asp Pro Glu Leu Glu Lys Lys Leu Lys Glu Asn Lys Ile Ser Leu Glu
            20                  25                  30
Ser Glu Tyr Glu Lys Ile Lys Asp Ser Lys Phe Asp Asp Trp Lys Asn
            35              40                  45
Ile Arg Gly Pro Arg Pro Trp Glu Asp Pro Asp Leu Leu Gln Gly Arg
    50              55                  60
Asn Pro Glu Ser Leu Lys Thr Lys Thr Thr
65                  70
```

<210> 190
<211> 267
<212> PRT

321

<213> Homo sapiens
<220>
<221> SIGNAL
<222> -21..-1
<400> 190

```
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
    -20                 -15                 -10
Ile Trp Thr Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr
-5                   1               5                       10
Leu His His Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr
            15              20                  25
Val Ala Pro Glu Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala
        30              35                  40
Val Leu Cys Ile Ala Thr Ile Tyr Val Arg Tyr Lys Gln Val His Ala
    45              50                  55
Leu Ser Pro Glu Glu Asn Val Ile Ile Lys Leu Asn Lys Ala Gly Leu
60              65              70                      75
Val Leu Gly Ile Leu Ser Cys Leu Gly Leu Ser Ile Val Ala Asn Phe
            80              85                  90
Gln Lys Thr Thr Leu Phe Ala Ala His Val Ser Gly Ala Val Leu Thr
        95              100                 105
Phe Gly Met Gly Ser Leu Tyr Met Phe Val Gln Thr Ile Leu Ser Tyr
    110             115                 120
Gln Met Gln Pro Lys Ile His Gly Lys Gln Val Phe Trp Ile Arg Leu
    125             130                 135
Leu Leu Val Ile Trp Cys Gly Val Ser Ala Leu Ser Met Leu Thr Cys
140             145                 150                     155
Ser Ser Val Leu His Ser Gly Asn Phe Gly Thr Asp Leu Glu Gln Lys
            160                 165                 170
Leu His Trp Asn Pro Glu Asp Lys Gly Tyr Ala Leu His Met Ile Thr
            175             180                 185
Thr Ala Ala Glu Trp Ser Met Ser Phe Ser Phe Phe Gly Phe Phe Leu
            190             195                 200
Thr Tyr Ile Arg Asp Phe Gln Lys Ile Ser Leu Arg Val Glu Ala Asn
        205             210                 215
Leu His Gly Leu Thr Leu Tyr Asp Thr Ala Pro Cys Pro Ile Asn Asn
220             225                 230                     235
Glu Arg Thr Arg Leu Leu Ser Arg Asp Ile Arg
            240                 245
```

<210> 191
<211> 108
<212> PRT
<213> Homo sapiens
<400> 191

```
Met Gly Cys Val Phe Gln Ser Thr Glu Asp Lys Cys Ile Phe Lys Ile
1               5               10              15
Asp Trp Thr Leu Ser Pro Gly Glu His Ala Lys Asp Glu Tyr Val Leu
            20              25              30
Tyr Tyr Tyr Ser Asn Leu Ser Val Pro Ile Gly Arg Phe Gln Asn Arg
        35              40              45
Val His Leu Met Gly Asp Ile Leu Cys Asn Asp Gly Ser Leu Leu Leu
    50              55              60
Gln Asp Val Gln Glu Ala Asp Gln Gly Thr Tyr Ile Cys Glu Ile Arg
65              70              75                      80
Leu Lys Gly Glu Ser Gln Val Phe Lys Lys Ala Val Val Leu His Val
            85              90                      95
Leu Pro Glu Glu Pro Lys Gly Thr Gln Met Leu Thr
```

193

```
<210> 192
<211> 69
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -46..-1
<400> 192
Met Ser Val Phe Trp Gly Phe Val Gly Phe Leu Val Pro Trp Phe Ile
    -45                 -40                 -35
Pro Lys Gly Pro Asn Arg Gly Val Ile Ile Thr Met Leu Val Thr Cys
-30                 -25                 -20                     -15
Ser Val Cys Cys Tyr Leu Phe Trp Leu Ile Ala Ile Leu Ala Gln Leu
                -10                 -5                      1
Asn Pro Leu Phe Gly Pro Gln Leu Lys Asn Glu Thr Ile Trp Tyr Leu
        5                   10                  15
Lys Tyr His Trp Pro
    20

<210> 193
<211> 251
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -28..-1
<400> 193
Met Trp Arg Leu Leu Ala Arg Ala Ser Ala Pro Leu Leu Arg Val Pro
            -25                 -20                 -15
Leu Ser Asp Ser Trp Ala Leu Leu Pro Ala Ser Ala Gly Val Lys Thr
            -10                 -5                  1
Leu Leu Pro Val Pro Ser Phe Glu Asp Val Ser Ile Pro Glu Lys Pro
5                   10                  15                  20
Lys Leu Arg Phe Ile Glu Arg Ala Pro Leu Val Pro Lys Val Arg Arg
                25                  30                  35
Glu Pro Lys Asn Leu Ser Asp Ile Arg Gly Pro Ser Thr Glu Ala Thr
                40                  45                  50
Glu Phe Thr Glu Gly Asn Phe Ala Ile Leu Ala Leu Gly Gly Gly Tyr
                55                  60                  65
Leu His Trp Gly His Phe Glu Met Met Arg Leu Thr Ile Asn Arg Ser
        70                  75                  80
Met Asp Pro Lys Asn Met Phe Ala Ile Trp Arg Val Pro Ala Pro Phe
85                  90                  95                      100
Lys Pro Ile Thr Arg Lys Ser Val Gly His Arg Met Gly Gly Gly Lys
                105                 110                 115
Gly Ala Ile Asp His Tyr Val Thr Pro Val Lys Ala Gly Arg Leu Val
                120                 125                 130
Val Glu Met Gly Gly Arg Cys Glu Phe Glu Glu Val Gln Gly Phe Leu
        135                 140                 145
Asp Gln Val Ala His Lys Leu Pro Phe Ala Ala Lys Ala Val Ser Arg
        150                 155                 160
Gly Thr Leu Glu Lys Met Arg Lys Asp Gln Glu Glu Arg Glu Arg Asn
165                 170                 175                     180
Asn Gln Asn Pro Trp Thr Phe Glu Arg Ile Ala Thr Ala Asn Met Leu
                185                 190                 195
Gly Ile Arg Lys Val Leu Ser Pro Tyr Asp Leu Thr His Lys Gly Lys
```

194   323

Tyr Trp Gly Lys Phe Tyr Met Pro Lys Arg Val
        215   220

<210> 194
<211> 99
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -48..-1
<400> 194
Met Asp Asn Val Gln Pro Lys Ile Lys His Arg Pro Phe Cys Phe Ser
            -45   -40   -35
Val Lys Gly His Val Lys Met Leu Arg Leu Asp Ile Ile Asn Ser Leu
        -30   -25   -20
Val Thr Thr Val Phe Met Leu Ile Val Ser Val Leu Ala Leu Ile Pro
    -15   -10   -5
Glu Thr Thr Thr Leu Thr Val Gly Gly Gly Val Phe Ala Leu Val Thr
1   5   10   15
Ala Val Cys Cys Leu Ala Asp Gly Ala Leu Ile Tyr Arg Lys Leu Leu
            20   25   30
Phe Asn Pro Ser Gly Pro Tyr Gln Lys Lys Pro Val His Glu Lys Lys
        35   40   45
Glu Val Leu
    50

<210> 195
<211> 81
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -31..-1
<400> 195
Met Ser Asn Thr His Thr Val Leu Val Ser Leu Pro His Pro His Pro
    -30   -25   -20
Ala Leu Thr Cys Cys His Leu Gly Leu Pro His Pro Val Arg Ala Pro
-15   -10   -5   1
Arg Pro Leu Pro Arg Val Glu Pro Trp Asp Pro Arg Trp Gln Asp Ser
        5   10   15
Glu Leu Arg Tyr Pro Gln Ala Met Asn Ser Phe Leu Asn Glu Arg Ser
        20   25   30
Ser Pro Cys Arg Thr Leu Arg Gln Glu Ala Ser Ala Asp Arg Cys Asp
    35   40   45
Leu
50

<210> 196
<211> 150
<212> PRT
<213> Homo sapiens
<400> 196
Met Lys Val His Met His Thr Lys Phe Cys Leu Ile Cys Leu Leu Thr
1   5   10   15
Phe Ile Phe His His Cys Asn His Cys His Glu Glu His Asp His Gly
        20   25   30
Pro Glu Ala Leu His Arg Gln His Arg Gly Met Thr Glu Leu Glu Pro

```
                35                      40                      45
Ser Lys Phe Ser Lys Gln Ala Ala Glu Asn Glu Lys Lys Tyr Tyr Ile
    50                      55                      60
Glu Lys Leu Phe Glu Arg Tyr Gly Glu Asn Gly Arg Leu Ser Phe Phe
65                      70                      75                      80
Gly Leu Glu Lys Leu Leu Thr Asn Leu Gly Leu Gly Glu Arg Lys Val
                85                      90                      95
Val Glu Ile Asn His Glu Asp Leu Gly His Asp His Val Ser His Leu
                100                     105                     110
Gly Ile Leu Ala Val Gln Glu Gly Lys His Phe His Ser His Asn His
                115                     120                     125
Gln His Ser His Asn His Leu Asn Ser Glu Asn Gln Thr Val Thr Ser
                130                     135                     140
Val Ser Thr Lys Lys Lys
145                     150


<210> 197
<211> 273
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -45..-1
<400> 197
Met Asn Trp Ser Ile Phe Glu Gly Leu Leu Ser Gly Val Asn Lys Tyr
-45                     -40                     -35                     -30
Ser Thr Ala Phe Gly Arg Ile Trp Leu Ser Leu Val Phe Ile Phe Arg
                -25                     -20                     -15
Val Leu Val Tyr Leu Val Thr Ala Glu Arg Val Trp Ser Asp Asp His
                -10                     -5                      1
Lys Asp Phe Asp Cys Asn Thr Arg Gln Pro Gly Cys Ser Asn Val Cys
    5                       10                      15
Phe Asp Glu Phe Phe Pro Val Ser His Val Arg Leu Trp Ala Leu Gln
20                      25                      30                      35
Leu Ile Leu Val Thr Cys Pro Ser Leu Leu Val Val Met His Val Ala
                40                      45                      50
Tyr Arg Glu Val Gln Glu Lys Arg His Arg Glu Ala His Gly Glu Asn
                55                      60                      65
Ser Gly Arg Leu Tyr Leu Asn Pro Gly Lys Lys Arg Gly Gly Leu Trp
                70                      75                      80
Trp Thr Tyr Val Cys Ser Leu Val Phe Lys Ala Ser Val Asp Ile Ala
                85                      90                      95
Phe Leu Tyr Val Phe His Ser Phe Tyr Pro Lys Tyr Ile Leu Pro Pro
100                     105                     110                     115
Val Val Lys Cys His Ala Asp Pro Cys Pro Asn Ile Val Asp Cys Phe
                120                     125                     130
Ile Ser Lys Pro Ser Glu Lys Asn Ile Phe Thr Leu Phe Met Val Ala
                135                     140                     145
Thr Ala Ala Ile Cys Ile Leu Leu Asn Leu Val Glu Leu Ile Tyr Leu
                150                     155                     160
Val Ser Lys Arg Cys His Glu Cys Leu Ala Ala Arg Lys Ala Gln Ala
                165                     170                     175
Met Cys Thr Gly His His Pro His Asp Thr Thr Ser Ser Cys Lys Gln
180                     185                     190                     195
Asp Asp Leu Leu Ser Gly Asp Leu Ile Phe Leu Gly Ser Asp Ser His
                200                     205                     210
Pro Pro Leu Leu Pro Asp Arg Pro Arg Asp His Val Lys Lys Thr Ile
                215                     220                     225
```

196

Leu

<210> 198
<211> 413
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -37..-1
<400> 198

```
Met Ala Ser Lys Ile Leu Leu Asn Val Gln Glu Glu Val Thr Cys Pro
        -35                   -30                   -25
Ile Cys Leu Glu Leu Leu Thr Glu Pro Leu Ser Leu Asp Cys Gly His
        -20                   -15                   -10
Ser Leu Cys Arg Ala Cys Ile Thr Val Ser Asn Lys Glu Ala Val Thr
-5                    1                     5                     10
Ser Met Gly Gly Lys Ser Ser Cys Pro Val Cys Gly Ile Ser Tyr Ser
                15                    20                    25
Phe Glu His Leu Gln Ala Asn Gln His Leu Ala Asn Ile Val Glu Arg
            30                    35                    40
Leu Lys Glu Val Lys Leu Ser Pro Asp Asn Gly Lys Lys Arg Asp Leu
    45                    50                    55
Cys Asp His His Gly Glu Lys Leu Leu Leu Phe Cys Lys Glu Asp Arg
60                    65                    70                    75
Lys Val Ile Cys Trp Leu Cys Glu Arg Ser Gln Glu His Arg Gly His
                80                    85                    90
His Thr Val Leu Thr Glu Glu Val Phe Lys Glu Cys Gln Glu Lys Leu
                95                    100                   105
Gln Ala Val Leu Lys Arg Leu Lys Lys Glu Glu Glu Glu Ala Glu Lys
            110                   115                   120
Leu Glu Ala Asp Ile Arg Glu Glu Lys Thr Ser Trp Lys Tyr Gln Val
    125                   130                   135
Gln Thr Glu Arg Gln Arg Ile Gln Thr Glu Phe Asp Gln Leu Arg Ser
140                   145                   150                   155
Ile Leu Asn Asn Glu Glu Gln Arg Glu Leu Gln Arg Leu Glu Glu Glu
                160                   165                   170
Glu Lys Lys Thr Leu Asp Lys Phe Ala Glu Ala Glu Asp Glu Leu Val
                175                   180                   185
Gln Gln Lys Gln Leu Val Arg Glu Leu Ile Ser Asp Val Glu Cys Arg
            190                   195                   200
Ser Gln Trp Ser Thr Met Glu Leu Leu Gln Asp Met Ser Gly Ile Met
    205                   210                   215
Lys Trp Ser Glu Ile Trp Arg Leu Lys Lys Pro Lys Met Val Ser Lys
220                   225                   230                   235
Lys Leu Lys Thr Val Phe His Ala Pro Asp Leu Ser Arg Met Leu Gln
                240                   245                   250
Met Phe Arg Glu Leu Thr Ala Val Arg Cys Tyr Trp Val Asp Val Thr
            255                   260                   265
Leu Asn Ser Val Asn Leu Asn Leu Val Leu Ser Glu Asp Gln
            270                   275                   280
Arg Gln Val Ile Ser Val Pro Ile Trp Pro Phe Gln Cys Tyr Asn Tyr
    285                   290                   295
Gly Val Leu Gly Ser Gln Tyr Phe Ser Ser Gly Lys His Tyr Trp Glu
300                   305                   310                   315
Val Asp Val Ser Lys Lys Thr Ala Trp Ile Leu Gly Val Tyr Cys Arg
                320                   325                   330
Thr Tyr Ser Arg His Met Lys Tyr Val Val Arg Arg Cys Ala Asn Arg
                335                   340                   345
```

197

```
Gln Asn Leu Tyr Thr Lys Tyr Arg Pro Leu Phe Gly Tyr Trp Val Ile
        350                 355                 360
Gly Leu Gln Asn Lys Cys Lys Tyr Gly Ala Lys Lys Lys
        365                 370                 375


<210> 199
<211> 393
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -19..-1
<400> 199
Met Arg Thr Leu Phe Asn Leu Leu Trp Leu Ala Leu Ala Cys Ser Pro
                -15                 -10                 -5
Val His Thr Thr Leu Ser Lys Ser Asp Ala Lys Lys Ala Ala Ser Lys
                1                   5                   10
Thr Leu Leu Glu Lys Ser Gln Phe Ser Asp Lys Pro Val Gln Asp Arg
        15                  20                  25
Gly Leu Val Val Thr Asp Leu Lys Ala Glu Ser Val Val Leu Glu His
30                  35                  40                  45
Arg Ser Tyr Cys Ser Ala Lys Ala Arg Asp Arg His Phe Ala Gly Asp
                50                  55                  60
Val Leu Gly Tyr Val Thr Pro Trp Asn Ser His Gly Tyr Asp Val Thr
                65                  70                  75
Lys Val Phe Gly Ser Lys Phe Thr Gln Ile Ser Pro Val Trp Leu Gln
                80                  85                  90
Leu Lys Arg Arg Gly Arg Glu Met Phe Glu Val Thr Gly Leu His Asp
        95                  100                 105
Val Asp Gln Gly Trp Met Arg Ala Val Arg Lys His Ala Lys Gly Leu
110                 115                 120                 125
His Ile Val Pro Arg Leu Leu Phe Glu Asp Trp Thr Tyr Asp Asp Phe
                130                 135                 140
Arg Asn Val Leu Asp Ser Glu Asp Glu Ile Glu Glu Leu Ser Lys Thr
                145                 150                 155
Val Val Gln Val Ala Lys Asn Gln His Phe Asp Gly Phe Val Val Glu
                160                 165                 170
Val Trp Asn Gln Leu Leu Ser Gln Lys Arg Val Gly Leu Ile His Met
        175                 180                 185
Leu Thr His Leu Ala Glu Ala Leu His Gln Ala Arg Leu Leu Ala Leu
190                 195                 200                 205
Leu Val Ile Pro Pro Ala Ile Thr Pro Gly Thr Asp Gln Leu Gly Met
                210                 215                 220
Phe Thr His Lys Glu Phe Glu Gln Leu Ala Pro Val Leu Asp Gly Phe
                225                 230                 235
Ser Leu Met Thr Tyr Asp Tyr Ser Thr Ala His Gln Pro Gly Pro Asn
        240                 245                 250
Ala Pro Leu Ser Trp Val Arg Ala Cys Val Gln Val Leu Asp Pro Lys
        255                 260                 265
Ser Lys Trp Arg Ser Lys Ile Leu Leu Gly Leu Asn Phe Tyr Gly Met
270                 275                 280                 285
Asp Tyr Ala Thr Ser Lys Asp Ala Arg Glu Pro Val Val Gly Ala Arg
                290                 295                 300
Tyr Ile Gln Thr Leu Lys Asp His Arg Pro Arg Met Val Trp Asp Ser
        305                 310                 315
Gln Ala Ser Glu His Phe Phe Glu Tyr Lys Lys Ser Arg Ser Gly Arg
        320                 325                 330
His Val Val Phe Tyr Pro Thr Leu Lys Ser Leu Gln Val Arg Leu Glu
```

198

```
                335                     340                     345
Leu Ala Arg Glu Leu Gly Val Gly Val Ser Ile Trp Glu Leu Gly Gln
350                     355                     360                     365
Gly Leu Asp Tyr Phe Tyr Asp Leu Leu
                    370


<210> 200
<211> 381
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -13..-1
<400> 200
Met Leu Leu Ser Ile Gly Met Leu Met Leu Ser Ala Thr Gln Val Tyr
                -10                     -5                      1
Thr Val Leu Thr Val Gln Leu Phe Ala Phe Leu Asn Pro Leu Pro Val
        5                       10                      15
Glu Ala Asp Ile Leu Ala Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe
20                      25                      30                      35
Asp Asp Leu Pro Ala Arg Phe Gly Tyr Arg Leu Pro Ala Glu Gly Leu
                    40                      45                      50
Lys Gly Phe Leu Ile Asn Ser Lys Pro Glu Asn Ala Cys Glu Pro Ile
                55                      60                      65
Val Pro Pro Pro Val Lys Asp Asn Ser Ser Gly Thr Phe Ile Val Leu
                70                      75                      80
Ile Arg Arg Leu Asp Cys Asn Phe Asp Ile Lys Val Leu Asn Ala Gln
        85                      90                      95
Arg Ala Gly Tyr Lys Ala Ala Ile Val His Asn Val Asp Ser Asp Asp
100                     105                     110                     115
Leu Ile Ser Met Gly Ser Asn Asp Ile Glu Val Leu Lys Lys Ile Asp
                    120                     125                     130
Ile Pro Ser Val Phe Ile Gly Glu Ser Ser Ala Ser Ser Leu Lys Asp
                135                     140                     145
Glu Phe Thr Tyr Glu Lys Gly Gly His Leu Ile Leu Val Pro Glu Phe
                150                     155                     160
Ser Leu Pro Leu Glu Tyr Tyr Leu Ile Pro Phe Leu Ile Ile Val Gly
        165                     170                     175
Ile Cys Leu Ile Leu Ile Val Ile Phe Met Ile Thr Lys Phe Val Gln
180                     185                     190                     195
Asp Arg His Arg Ala Arg Arg Asn Arg Leu Arg Lys Asp Gln Leu Lys
                    200                     205                     210
Lys Leu Pro Val His Lys Phe Lys Lys Gly Asp Glu Tyr Asp Val Cys
                215                     220                     225
Ala Ile Cys Leu Asp Glu Tyr Glu Asp Gly Asp Lys Leu Arg Ile Leu
                230                     235                     240
Pro Cys Ser His Ala Tyr His Cys Lys Cys Val Asp Pro Trp Leu Thr
                245                     250                     255
Lys Thr Lys Lys Thr Cys Pro Val Cys Arg Gln Lys Val Val Pro Ser
260                     265                     270                     275
Gln Gly Asp Ser Asp Ser Asp Thr Asp Ser Ser Gln Glu Glu Asn Glu
                    280                     285                     290
Val Thr Glu His Thr Pro Leu Leu Arg Pro Leu Ala Ser Val Ser Ala
                295                     300                     305
Gln Ser Phe Gly Ala Leu Ser Glu Ser Arg Ser His Gln Asn Met Thr
                310                     315                     320
Glu Ser Ser Asp Tyr Glu Glu Asp Asp Asn Glu Asp Thr Asp Ser Ser
                325                     330                     335
```

199

Asp Ala Glu Asn Glu Ile Asn Glu His Asp Val Val Val Gln Leu Gln
340   345   350   355
Pro Asn Gly Glu Arg Asp Tyr Asn Ile Ala Asn Thr Val
                360   365

<210> 201
<211> 291
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 201
Met Asp Ser Arg Val Ser Ser Pro Glu Lys Gln Asp Lys Glu Asn Phe
        -40   -35   -30
Val Gly Val Asn Asn Lys Arg Leu Gly Val Cys Gly Trp Ile Leu Phe
        -25   -20   -15
Ser Leu Ser Phe Leu Leu Val Ile Ile Thr Phe Pro Ile Ser Ile Trp
-10   -5   1   5
Met Cys Leu Lys Ile Ile Arg Glu Tyr Glu Arg Ala Val Val Phe Arg
            10   15   20
Leu Gly Arg Ile Gln Ala Asp Lys Ala Lys Gly Pro Gly Leu Ile Leu
            25   30   35
Val Leu Pro Cys Ile Asp Val Phe Val Lys Val Asp Leu Arg Thr Val
        40   45   50
Thr Cys Asn Ile Pro Pro Gln Glu Ile Leu Thr Arg Asp Ser Val Thr
55   60   65   70
Thr Gln Val Asp Gly Val Val Tyr Tyr Arg Ile Tyr Ser Ala Val Ser
                75   80   85
Ala Val Ala Asn Val Asn Asp Val His Gln Ala Thr Phe Leu Leu Ala
            90   95   100
Gln Thr Thr Leu Arg Asn Val Leu Gly Thr Gln Thr Leu Ser Gln Ile
            105   110   115
Leu Ala Gly Arg Glu Glu Ile Ala His Ser Ile Gln Thr Leu Leu Asp
            120   125   130
Asp Ala Thr Glu Leu Trp Gly Ile Arg Val Ala Arg Val Glu Ile Lys
135   140   145   150
Asp Val Arg Ile Pro Val Gln Leu Gln Arg Ser Met Ala Ala Glu Ala
                155   160   165
Glu Ala Thr Arg Glu Ala Arg Ala Lys Val Leu Ala Ala Glu Gly Glu
            170   175   180
Met Ser Ala Ser Lys Ser Leu Lys Ser Ala Ser Met Val Leu Ala Glu
            185   190   195
Ser Pro Ile Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu Ser Thr Val
        200   205   210
Ala Thr Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro Met Asn Ile
215   220   225   230
Leu Glu Gly Ile Gly Gly Val Ser Tyr Asp Asn His Lys Lys Leu Pro
                235   240   245
Asn Lys Ala

<210> 202
<211> 92
<212> PRT
<213> Homo sapiens
<400> 202
Met Pro Pro Arg Asn Leu Leu Glu Leu Leu Ile Asn Ile Lys Ala Gly
1   5   10   15

```
Thr Tyr Leu Pro Gln Ser Tyr Leu Ile His Glu His Met Val Ile Thr
            20              25              30
Asp Arg Ile Glu Asn Ile Asp His Leu Gly Phe Phe Ile Tyr Arg Leu
        35              40              45
Cys His Asp Lys Glu Thr Tyr Lys Leu Gln Arg Arg Glu Thr Ile Lys
    50          .       55              60
Gly Ile Gln Lys Arg Glu Ala Ser Asn Cys Phe Ala Ile Arg His Phe
65              70              75              80
Glu Asn Lys Phe Ala Val Glu Thr Leu Ile Cys Ser
            85              90

<210> 203
<211> 127
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -63..-1
<400> 203
Met Ser Ala Ala Gly Ala Arg Gly Leu Arg Ala Thr Tyr His Arg Leu
            -60             -55             -50
Pro Asp Lys Val Glu Leu Met Leu Pro Glu Lys Leu Arg Pro Leu Tyr
            -45             -40             -35
Asn His Pro Ala Gly Pro Arg Thr Val Phe Phe Trp Ala Pro Ile Met
        -30             -25             -20
Lys Trp Gly Leu Val Cys Ala Gly Leu Ala Asp Met Ala Arg Pro Ala
-15             -10             -5                      1
Glu Lys Leu Ser Thr Ala Gln Ser Ala Val Leu Met Ala Thr Gly Phe
            5               10              15
Ile Trp Ser Arg Tyr Ser Leu Val Ile Ile Pro Lys Asn Trp Ser Leu
            20              25              30
Phe Ala Val Asn Phe Phe Val Gly Ala Ala Gly Ala Ser Gln Leu Phe
        35              40              45
Arg Ile Trp Arg Tyr Asn Gln Glu Leu Lys Ala Lys Ala His Lys
50              55              60

<210> 204
<211> 84
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -20..-1
<400> 204
Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly Ala Leu Leu Gly
-20             -15             -10             -5
Thr Ala Trp Ala Arg Arg Ser Gln Asp Leu His Cys Gly Ala Cys Arg
                1               5               10
Ala Leu Val Asp Glu Leu Glu Trp Glu Ile Ala Gln Val Asp Pro Lys
        15              20              25
Lys Thr Ile Gln Met Gly Ser Phe Arg Ile Asn Pro Asp Gly Ser Gln
    30              35              40
Ser Val Val Glu Val Thr Val Thr Val Pro Pro Asn Lys Val Ala His
45              50              55              60
Ser Gly Phe Gly

<210> 205
<211> 182
```

201

<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -20..-1
<400> 205
Met Lys Gly Trp Gly Trp Leu Ala Leu Leu Leu Gly Ala Leu Leu Gly
-20   -15   -10   -5
Thr Ala Trp Ala Arg Arg Ser Gln Asp Leu His Cys Gly Ala Cys Arg
              1   5   10
Ala Leu Val Asp Glu Leu Glu Trp Glu Ile Ala Gln Val Asp Pro Lys
          15   20   25
Lys Thr Ile Gln Met Gly Ser Phe Arg Ile Asn Pro Asp Gly Ser Gln
      30   35   40
Ser Val Val Glu Val Pro Tyr Ala Arg Ser Glu Ala His Leu Thr Glu
45   50   55   60
Leu Leu Glu Glu Ile Cys Asp Arg Met Lys Glu Tyr Gly Glu Gln Ile
              65   70   75
Asp Pro Ser Thr His Arg Lys Asn Tyr Val Arg Val Val Gly Arg Asn
          80   85   90
Gly Glu Ser Ser Glu Leu Asp Leu Gln Gly Ile Arg Ile Asp Ser Asp
          95   100   105
Ile Ser Gly Thr Leu Lys Phe Ala Cys Gly Ser Ile Val Glu Glu Tyr
      110   115   120
Glu Asp Glu Leu Ile Glu Phe Phe Ser Arg Glu Ala Asp Asn Val Lys
125   130   135   140
Asp Lys Leu Cys Ser Lys Arg Thr Asp Leu Cys Asp His Ala Leu His
              145   150   155
Ile Ser His Asp Glu Leu
              160

<210> 206
<211> 71
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -25..-1
<400> 206
Met Pro Ala Gly Val Pro Met Ser Thr Tyr Leu Lys Met Phe Ala Ala
-25   -20   -15   -10
Ser Leu Leu Ala Met Cys Ala Gly Ala Glu Val Val His Arg Tyr Tyr
              -5   1   5
Arg Pro Asp Leu Thr Ile Pro Glu Ile Pro Pro Lys Arg Gly Glu Leu
          10   15   20
Lys Thr Glu Leu Leu Gly Leu Lys Glu Arg Lys His Lys Pro Gln Val
      25   30   35
Ser Gln Gln Glu Glu Leu Lys
40   45

<210> 207
<211> 73
<212> PRT
<213> Homo sapiens
<400> 207
Met Arg Ile Arg Met Thr Asp Gly Arg Thr Leu Val Gly Cys Phe Leu
1   5   10   15
Cys Thr Asp Arg Asp Cys Asn Val Ile Leu Gly Ser Ala Gln Glu Phe

202 331

```
                    20                      25                      30
Leu Lys Pro Ser Asp Ser Phe Ser Ala Gly Glu Pro Arg Val Leu Gly
          35                      40                      45
Leu Ala Met Val Pro Gly His His Ile Val Ser Ile Glu Val Gln Arg
      50                      55                      60
Glu Ser Leu Thr Gly Pro Pro Tyr Leu
65                      70
```

<210> 208
<211> 169
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -150..-1
<220>
<221> UNSURE
<222> -67
<223> Xaa = any one of the twenty amino acids
<400> 208

```
Met Ala Glu Thr Lys Asp Thr Ala Gln Met Leu Val Thr Phe Lys Asp
-150                    -145                    -140                    -135
Val Ala Val Thr Phe Thr Arg Glu Glu Trp Arg Gln Leu Asp Leu Ala
                  -130                    -125                    -120
Gln Arg Thr Leu Tyr Arg Glu Gly Ile Gly Phe Pro Lys Pro Glu Leu
                  -115                    -110                    -105
Val His Leu Leu Glu His Gly Gln Glu Leu Trp Ile Val Lys Arg Gly
              -100                    -95                      -90
Leu Ser His Ala Thr Cys Ala Glu Phe His Ser Cys Cys Pro Gly Trp
          -85                      -80                      -75
Ser Ala Val Xaa Arg His Leu Ser Ser Leu Gln Leu Leu Pro Pro Glu
-70                      -65                      -60                      -55
Phe Lys Gly Phe Ser Cys Leu Ser Leu Pro Ser Ser Trp Asp Tyr Arg
                  -50                      -45                      -40
Arg Pro Pro Pro Cys Pro Ala Gly Phe Phe Val Phe Leu Val Glu Thr
                  -35                      -30                      -25
Gly Leu His His Val Gly Gln Ala Gly Leu Glu Leu Leu Thr Ser Cys
          -20                      -15                      -10
Ser Pro Pro Ala Ser Ala Ser Gln Ser Ala Ala Ile Thr Gly Val Ser
      -5                      1                       5                       10
His Arg Ala Arg Gln Arg Lys Thr Ala
                  15
```

<210> 209
<211> 76
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<400> 209

```
Met Glu Leu Ile Ser Pro Thr Val Ile Ile Ile Leu Gly Cys Leu Ala
          -20                      -15                      -10
Leu Phe Leu Leu Leu Gln Arg Lys Asn Leu Arg Arg Pro Pro Cys Ile
      -5                      1                       5                       10
Lys Gly Trp Ile Pro Trp Ile Gly Val Gly Phe Glu Phe Gly Lys Ala
                  15                      20                      25
Pro Leu Glu Phe Ile Glu Lys Ala Arg Ile Lys Val Cys Gly Arg Gly
```

```
                     30                        35                        40
Arg Arg Gly Leu Gln Arg Arg Gln Cys Phe Leu Phe
            45                        50


<210> 210
<211> 95
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -54..-1
<400> 210
Met Ala Glu Thr Lys Asp Ala Ala Gln Met Leu Val Thr Phe Lys Asp
                    -50                       -45                       -40
Val Ala Val Thr Phe Thr Arg Glu Glu Trp Arg Gln Leu Asp Leu Ala
                    -35                       -30                       -25
Gln Arg Thr Leu Tyr Arg Glu Val Met Leu Glu Thr Cys Gly Leu Leu
                -20                       -15                       -10
Val Ser Leu Val Glu Ser Ile Trp Leu His Ile Thr Glu Asn Gln Ile
            -5                        1                         5                         10
Lys Leu Ala Ser Pro Gly Arg Lys Phe Thr Asn Ser Pro Asp Glu Lys
                    15                        20                        25
Pro Glu Val Trp Leu Ala Pro Gly Leu Phe Gly Ala Ala Ala Gln
                30                        35                        40


<210> 211
<211> 92
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<400> 211
Met Glu Leu Ile Ser Pro Thr Val Ile Ile Ile Leu Gly Cys Leu Ala
                -20                       -15                       -10
Leu Phe Leu Leu Leu Gln Arg Lys Asn Leu Arg Arg Pro Pro Cys Ile
            -5                        1                         5                         10
Lys Gly Trp Ile Pro Trp Ile Gly Val Gly Phe Glu Phe Gly Lys Ala
                    15                        20                        25
Pro Leu Glu Phe Ile Glu Lys Ala Arg Ile Lys Tyr Gly Pro Ile Phe
                30                        35                        40
Thr Val Phe Ala Met Gly Asn Arg Met Thr Phe Val Thr Glu Glu Glu
                45                        50                        55
Gly Ile Asn Val Phe Leu Lys Ser Lys Lys Lys Lys
            60                        65                        70


<210> 212
<211> 89
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -16..-1
<400> 212
Met Ile Ile Ser Leu Phe Ile Tyr Ile Phe Leu Thr Cys Ser Asn Thr
            -15                       -10                       -5
Ser Pro Ser Tyr Gln Gly Thr Gln Leu Gly Leu Gly Leu Pro Ser Ala
1                         5                                   10                        15
```

```
Gln Trp Trp Pro Leu Thr Gly Arg Arg Met Gln Cys Cys Arg Leu Phe
            20                  25              30
Cys Phe Leu Leu Gln Asn Cys Leu Phe Pro Phe Pro Leu His Leu Ile
        35                  40              45
Gln His Asp Pro Cys Glu Leu Val Leu Thr Ile Ser Trp Asp Trp Ala
    50                  55                  60
Glu Ala Gly Ala Ser Leu Tyr Ser Pro
65                  70
```

<210> 213
<211> 109
<212> PRT
<213> Homo sapiens
<400> 213

```
Met Lys Val Asp Lys Asp Arg Gln Met Val Val Leu Glu Glu Glu Phe
1               5                   10                  15
Arg Asn Ile Ser Pro Glu Glu Leu Lys Met Glu Leu Pro Glu Arg Gln
            20                  25                  30
Pro Arg Phe Val Val Tyr Ser Tyr Lys Tyr Val Arg Asp Asp Gly Arg
        35                  40                  45
Val Ser Tyr Pro Leu Cys Phe Ile Phe Ser Ser Pro Val Gly Cys Lys
    50                  55                  60
Pro Glu Gln Gln Met Met Tyr Ala Gly Ser Lys Asn Arg Leu Val Gln
65                  70                  75                  80
Thr Ala Glu Leu Thr Lys Val Phe Glu Ile Arg Thr Thr Asp Asp Leu
                85                  90                  95
Thr Glu Ala Trp Leu Gln Glu Lys Leu Ser Phe Phe Arg
            100                 105
```

<210> 214
<211> 114
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -103..-1
<400> 214

```
Met Val Ile Arg Val Tyr Ile Ala Ser Ser Ser Gly Ser Thr Ala Ile
            -100                -95                 -90
Lys Lys Lys Gln Gln Asp Val Leu Gly Phe Leu Glu Ala Asn Lys Ile
        -85                 -80                 -75
Gly Phe Glu Glu Lys Asp Ile Ala Ala Asn Glu Glu Asn Arg Lys Trp
        -70                 -65                 -60
Met Arg Glu Asn Val Pro Glu Asn Ser Arg Pro Ala Thr Gly Asn Pro
-55                 -50                 -45                 -40
Leu Pro Pro Gln Ile Phe Asn Glu Ser Gln Tyr Arg Gly Asp Tyr Asp
            -35                 -30                 -25
Ala Phe Phe Glu Ala Arg Glu Asn Asn Ala Val Tyr Ala Phe Leu Gly
        -20                 -15                 -10
Leu Thr Ala Pro Ser Gly Ser Lys Glu Ala Glu Val Gln Ala Lys Gln
        -5                  1                   5
Gln Ala
10
```

<210> 215
<211> 124
<212> PRT
<213> Homo sapiens

```
<220>
<221> SIGNAL
<222> -97..-1
<400> 215
Met Ala Asp Asp Leu Lys Arg Phe Leu Tyr Lys Lys Leu Pro Ser Val
        -95                 -90                 -85
Glu Gly Leu His Ala Ile Val Val Ser Asp Arg Asp Gly Val Pro Val
    -80                 -75      .                -70
Ile Lys Val Ala Asn Asp Asn Ala Pro Glu His Ala Leu Arg Pro Gly
-65                 -60                 -55                     -50
Phe Leu Ser Thr Phe Ala Leu Ala Thr Asp Gln Gly Ser Lys Leu Gly
                -45                 -40                 -35
Leu Ser Lys Asn Lys Ser Ile Ile Cys Tyr Tyr Asn Thr Tyr Gln Val
            -30                 -25              .  -20
Val Gln Phe Asn Arg Leu Pro Leu Val Val Ser Phe Ile Ala Ser Ser
    .      -15                 -10                 -5
Ser Ala Asn Thr Gly Leu Ile Val Ser Leu Glu Lys Glu Leu Ala Pro
    1               5                   10                      15
Leu Phe Glu Glu Leu Arg Gln Val Val Glu Val Ser
                20              25


<210> 216
<211> 93
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -22..-1
<400> 216
Met Lys Pro Val Leu Pro Leu Gln Phe Leu Val Val Phe Cys Leu Ala
        -20                 -15                 -10
Leu Gln Leu Val Pro Gly Ser Pro Lys Gln Arg Val Leu Lys Tyr Ile
    -5              .1          .       5                   10
Leu Glu Pro Pro Pro Cys Ile Ser Ala Pro Glu Asn Cys Thr His Leu
                15              .    20                  25
Cys Thr Met Gln Glu Asp Cys Glu Lys Gly Phe Gln Cys Cys Ser Ser
            30                  35                  40
Phe Cys Gly Ile Val Cys Ser Ser Glu Thr Phe Gln Lys Arg Asn Arg
            45                  50                  55
Ile Lys His Lys Gly Ser Glu Val Ile Met Pro Ala Asn
    60                  65                  70


<210> 217
<211> 207
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -42..-1
<400> 217
Met His Ile Leu Gln Leu Leu Thr Thr Val Asp Asp Gly Ile Gln Ala
        -40                 -35                 -30
Ile Val His Cys Pro Asp Thr Gly Lys Asp Ile Trp Asn Leu Leu Phe
    -25                 -20                 -15
Asp Leu Val Cys His Glu Phe Cys Gln Ser Asp Asp Pro Pro Ile Ile
-10         .           -5                  1               5
Leu Gln Glu Gln Lys Thr Val Leu Ala Ser Val Phe Ser Val Leu Ser
            10                  15                  20
```

```
Ala Ile Tyr Ala Ser Gln Thr Glu Gln Glu Tyr Leu Lys Ile Glu Lys
        25                  30                  35
Val Asp Leu Pro Leu Ile Asp Ser Leu Ile Arg Val Leu Gln Asn Met
        40                  45                  50
Glu Gln Cys Gln Lys Lys Pro Glu Asn Ser Ala Glu Ser Asn Thr Glu
55                      60                  65                  70
Glu Thr Lys Arg Thr Asp Leu Thr Gln Asp Asp Phe His Leu Lys Ile
                75                  80                  85
Leu Lys Asp Ile Leu Cys Glu Phe Leu Ser Asn Ile Phe Gln Ala Leu
                90                  95                  100
Thr Lys Glu Thr Val Ala Gln Gly Val Lys Glu Gly Gln Leu Ser Lys
                105                 110                 115
Gln Lys Cys Ser Ser Ala Phe Gln Asn Leu Leu Pro Phe Tyr Ser Pro
        120                 125                 130
Val Val Glu Asp Phe Ile Lys Ile Leu Arg Glu Val Asp Lys Ala Leu
135                 140                 145                 150
Ala Asp Asp Leu Glu Lys Asn Phe Pro Ser Leu Lys Val Gln Thr
                155                 160                 165


<210> 218
<211> 59
<212> PRT
<213> Homo sapiens
<400> 218
Met Pro His Ser Lys Pro Leu Asp Trp Gly Leu Ser Ser Val Ala Glu
1                   5                   10                  15
Cys Pro Ala Glu Leu Phe Pro Ser Thr Gly Gly Leu Ala Gly Lys Gly
                20                  25                  30
Pro Gly Leu Asp Ile Leu Arg Cys Val Leu Ser Pro Trp Ala Ser His
                35                  40                  45
Phe Pro Ser Leu Ser Leu Gly Val Phe Asn Leu
        50                  55


<210> 219
<211> 56
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -27..-1
<400> 219
Met Asn Arg Val Pro Ala Asp Ser Pro Asn Met Cys Leu Ile Cys Leu
        -25                 -20                 -15
Leu Ser Tyr Ile Ala Leu Gly Ala Ile His Ala Lys Ile Cys Arg Arg
        -10                 -5                  1                   5
Ala Phe Gln Glu Glu Gly Arg Ala Asn Ala Lys Thr Gly Val Arg Ala
                10                  15                  20
Trp Cys Ile Gln Pro Trp Ala Lys
                25


<210> 220
<211> 162
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -94..-1
<400> 220
```

```
Met Leu Gln Thr Ser Asn Tyr Ser Leu Val Leu Ser Leu Gln Phe Leu
            -90              -85              -80
Leu Leu Ser Tyr Asp Leu Phe Val Asn Ser Phe Ser Glu Leu Leu Gln
            -75              -70              -65
Lys Thr Pro Val Ile Gln Leu Val Leu Phe Ile Ile Gln Asp Ile Ala
            -60              -55              -50
Val Leu Phe Asn Ile Ile Ile Ile Phe Leu Met Phe Phe Asn Thr Phe
    -45              -40                  -35
Val Phe Gln Ala Gly Leu Val Asn Leu Leu Phe His Lys Phe Lys Gly
-30                  -25                  -20                  -15
Thr Ile Ile Leu Thr Ala Val Tyr Phe Ala Leu Ser Ile Ser Leu His
            -10               -5                       1
Val Trp Val Met Asn Leu Arg Trp Lys Asn Ser Asn Ser Phe Ile Trp
      5                  10                  15
Thr Asp Gly Leu Gln Met Leu Phe Val Phe Gln Arg Leu Ala Ala Val
    20                  25                  30
Leu Tyr Cys Tyr Phe Tyr Lys Arg Thr Ala Val Arg Leu Gly Asp Pro
35                  40                  45                       50
His Phe Tyr Gln Asp Ser Leu Trp Leu Arg Lys Glu Phe Met Gln Val
            55                  60                  65
Arg Arg

<210> 221
<211> 154
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -68..-1
<400> 221
Met Ala Ser Ala Ser Ala Arg Gly Asn Gln Asp Lys Asp Ala His Phe
            -65              -60              -55
Pro Pro Pro Ser Lys Gln Ser Leu Leu Phe Cys Pro Lys Ser Lys Leu
            -50              -45              -40
His Ile His Arg Ala Glu Ile Ser Lys Ile Met Arg Glu Cys Gln Glu
    -35              -30                  -25
Glu Ser Phe Trp Lys Arg Ala Leu Pro Phe Ser Leu Val Ser Met Leu
-20                  -15                  -10                   -5
Val Thr Gln Gly Leu Val Tyr Gln Gly Tyr Leu Ala Ala Asn Ser Arg
                  1               5                      10
Phe Gly Ser Leu Pro Lys Val Ala Leu Ala Gly Leu Leu Gly Phe Gly
    15                  20                  25
Leu Gly Lys Val Ser Tyr Ile Gly Val Cys Gln Ser Lys Phe His Phe
    30                  35                  40
Phe Glu Asp Gln Leu Arg Gly Ala Gly Phe Gly Pro Gln His Asn Arg
45                  50                  55                       60
His Cys Leu Leu Thr Cys Glu Glu Cys Lys Ile Lys His Gly Leu Ser
            65                  70                       75
Glu Lys Gly Asp Ser Gln Pro Ser Ala Ser
            80                  85

<210> 222
<211> 99
<212> PRT
<213> Homo sapiens
<400> 222
Met Lys Val Glu Glu Glu His Thr Asn Ala Ile Gly Thr Leu His Gly
1                 5                  10                  15
```

```
Gly Leu Thr Ala Thr Leu Val Asp Asn Ile Ser Thr Met Ala Leu Leu
            20                  25                  30
Cys Thr Glu Arg Gly Ala Pro Gly Val Ser Val Asp Met Asn Ile Thr
            35                  40                  45
Tyr Met Ser Pro Ala Lys Leu Gly Glu Asp Ile Val Ile Thr Ala His
    50                  55                  60
Val Leu Lys Gln Gly Lys Thr Leu Ala Phe Thr Ser Val Gly Leu Thr
65                  70                  75                  80
Asn Lys Ala Thr Gly Lys Leu Ile Ala Gln Gly Arg His Thr Lys His
                85                  90                  95
Leu Gly Asn

<210> 223
<211> 43
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -24..-1
<400> 223
Met Gln Cys Phe Ser Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu
                -20                 -15                 -10
Leu Ile Phe Leu Cys Gly Phe Thr Asn Tyr Thr Asp Phe Glu Asp Ser
                -5                  1                   5
Pro Tyr Phe Lys Met His Lys Pro Val Thr Met
    10                  15

<210> 224
<211> 69
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -21..-1
<400> 224
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
    -20                 -15                 -10
Ile Trp Thr Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr
-5                  1                   5                   10
Leu His His Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr
            15                  20                  25
Val Ala Pro Glu Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala
        30                  35                  40
Val Leu Cys Gln Lys
    45

<210> 225
<211> 78
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -18..-1
<400> 225
Met Ser Pro Gly Ser Ala Leu Ala Leu Leu Trp Ser Leu Pro Ala Ser
            -15                 -10                 -5
Asp Leu Gly Arg Ser Val Ile Ala Gly Leu Trp Pro His Thr Gly Val
        1                   5                   10
```

```
Leu Ile His Leu Glu Thr Ser Gln Ser Phe Leu Gln Gly Gln Leu Thr
15              20              25              30
Lys Ser Ile Phe Pro Leu Cys Cys Thr Ser Leu Phe Cys Val Cys Val
            35              40                      45
Val Thr Val Gly Gly Gly Arg Val Gly Ser Thr Phe Val Ala
            50              55                      60
```

<210> 226
<211> 80
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -47..-1
<400> 226

```
Met Arg Leu Pro Pro Ala Leu Pro Ser Gly Tyr Thr Asp Ser Thr Ala
        -45                 -40                 -35
Leu Glu Gly Leu Val Tyr Tyr Leu Asn Gln Lys Leu Leu Phe Ser Ser
    -30                 -25                 -20
Pro Ala Ser Ala Leu Leu Phe Phe Ala Arg Pro Cys Val Phe Cys Phe
-15                 -10                 -5                      1
Lys Ala Ser Lys Met Gly Pro Gln Phe Glu Asn Tyr Pro Thr Phe Pro
            5                   10                  15
Thr Tyr Ser Pro Leu Pro Ile Ile Pro Phe Gln Leu His Gly Arg Phe
        20                  25                  30
```

<210> 227
<211> 241
<212> PRT
<213> Homo sapiens
<220>
<221> SIGNAL
<222> -103..-1
<400> 227

```
Met Trp Leu Asp Pro Val Phe Pro Leu Phe Pro Val Gly Asp His Tyr
        -100                -95                 -90
Leu Pro His Leu His Met Asp Val Leu Glu Gly Leu Ile Leu Val Leu
        -85                 -80                 -75
Pro Cys Ile Asp Val Phe Val Lys Val Asp Leu Arg Thr Val Thr Cys
        -70                 -65                 -60
Asn Ile Pro Pro Gln Glu Ile Leu Thr Arg Asp Ser Val Thr Thr Gln
-55                 -50                 -45                 -40
Val Asp Gly Val Val Tyr Tyr Arg Ile Tyr Ser Ala Val Ser Ala Val
            -35                 -30                 -25
Ala Asn Val Asn Asp Val His Gln Ala Thr Phe Leu Leu Ala Gln Thr
            -20                 -15                 -10
Thr Leu Arg Asn Val Leu Gly Thr Gln Thr Leu Ser Gln Ile Leu Ala
        -5                  1                   5
Gly Arg Glu Glu Ile Ala His Ser Ile Gln Thr Leu Leu Asp Asp Ala
10                  15                  20                  25
Thr Glu Leu Trp Gly Ile Arg Val Ala Arg Val Glu Ile Lys Asp Val
            30                  35                  40
Arg Ile Pro Val Gln Leu Gln Arg Ser Met Ala Ala Glu Ala Glu Ala
            45                  50                  55
Thr Arg Glu Ala Arg Ala Lys Val Leu Ala Ala Glu Gly Glu Met Asn
            60                  65                  70
Ala Ser Lys Ser Leu Lys Ser Ala Ser Met Val Leu Ala Glu Ser Pro
    75                  80                  85
```

210

```
Ile Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu Ser Thr Val Ala Thr
90              95                  100                 105
Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro Met Asn Ile Leu Glu
                110             115                 120
Gly Ile Gly Gly Val Ser Tyr Asp Asn His Lys Lys Leu Pro Asn Lys
            125             130                 135
Ala

<210>   228
<211>   560
<212>   DNA
<213>   Homo sapiens
<220>
<221>   CDS
<222>   (11)..(439)
<220>
<221>   polyA_site
<222>   (547)..(560)
<220>
<221>   polyA_signal
<222>   (530)..(535)
<400>   228
cagaacaatc atg tct gac tcc ctg gtg gtg tgc gag gta gac cca gag          49
           Met Ser Asp Ser Leu Val Val Cys Glu Val Asp Pro Glu
           1               5                   10
cta aca gaa aag ctg agg aaa ttc cgc ttc cga aaa gag aca gac aat         97
Leu Thr Glu Lys Leu Arg Lys Phe Arg Phe Arg Lys Glu Thr Asp Asn
    15                  20                  25
gca gcc atc ata atg aag gtg gac aaa gac cgg cag atg gtg gtg ctg        145
Ala Ala Ile Ile Met Lys Val Asp Lys Asp Arg Gln Met Val Val Leu
30                  35                  40                  45
gag gaa gaa ttt cgg aac att tcc cca gag gag ctc aaa atg gag ttg        193
Glu Glu Glu Phe Arg Asn Ile Ser Pro Glu Glu Leu Lys Met Glu Leu
                50                  55                  60
ccg gag aga cag ccc agg ttc gtg gtt tac agc tac aag tac gtg cgt        241
Pro Glu Arg Gln Pro Arg Phe Val Val Tyr Ser Tyr Lys Tyr Val Arg
                65                  70                  75
gac gat ggc cga gtg tcc tac cct ttg tgt ttc atc ttc tcc agc cct        289
Asp Asp Gly Arg Val Ser Tyr Pro Leu Cys Phe Ile Phe Ser Ser Pro
            80                  85                  90
gtg ggc tgc aag ccg gaa caa cag atg atg tat gca ggg agt aaa aac        337
Val Gly Cys Lys Pro Glu Gln Gln Met Met Tyr Ala Gly Ser Lys Asn
    95                  100                 105
agg ctg gtg cag aca gca gag ctc aca aag gtg ttc gaa atc cgc acc        385
Arg Leu Val Gln Thr Ala Glu Leu Thr Lys Val Phe Glu Ile Arg Thr
110                 115                 120                 125
act gat gac ctc act gag gcc tgg ctc caa gaa aag ttg tct ttc ttt        433
Thr Asp Asp Leu Thr Glu Ala Trp Leu Gln Glu Lys Leu Ser Phe Phe
                130                 135                 140
cgt tga tctctgggct ggggactgaa ttcctgatgt ctgagtcctc aaggtgactg        489
Arg

gggacttgga acccctagga cctgaacaac caagacttta aataaatttt aaaatgcaaa        549
aaaaaaaaaa a                                                            560

<210>   229
<211>   142
<212>   PRT
```
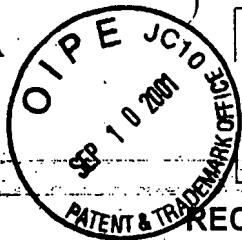
<213> Homo sapiens
<400> 229

```
Met Ser Asp Ser Leu Val Val Cys Glu Val Asp Pro Glu Leu Thr Glu
1               5                   10                  15
Lys Leu Arg Lys Phe Arg Phe Arg Lys Glu Thr Asp Asn Ala Ala Ile
            20                  25                  30
Ile Met Lys Val Asp Lys Asp Arg Gln Met Val Val Leu Glu Glu Glu
            35                  40                  45
Phe Arg Asn Ile Ser Pro Glu Glu Leu Lys Met Glu Leu Pro Glu Arg
        50                  55                  60
Gln Pro Arg Phe Val Val Tyr Ser Tyr Lys Tyr Val Arg Asp Asp Gly
65                  70                  75                  80
Arg Val Ser Tyr Pro Leu Cys Phe Ile Phe Ser Ser Pro Val Gly Cys
                85                  90                  95
Lys Pro Glu Gln Gln Met Met Tyr Ala Gly Ser Lys Asn Arg Leu Val
            100                 105                 110
Gln Thr Ala Glu Leu Thr Lys Val Phe Glu Ile Arg Thr Thr Asp Asp
            115                 120                 125
Leu Thr Glu Ala Trp Leu Gln Glu Lys Leu Ser Phe Phe Arg
    130                 135                 140
```

FORM PTO-1619A
Expires 06/30/99
OMB 0651-0027

U.S. Department of Commerce
Patent and Trademark Office
PATENT

RECORDATION FORM COVER SHEET
PATENTS ONLY

TO: The Commissioner of Patents and Trademarks: Please record the attached original document(s) or copy(ies).

Submission Type	Conveyance Type
[x] New	[x] Assignment [ ] Security Agreement
[ ] Resubmission (Non-Recordation) Document ID#	[ ] License [ ] Change of Name
[ ] Correction of PTO Error Reel #_____ Frame #_____	[ ] Merger [ ] Other _____
[ ] Corrective Document Reel #_____ Frame #_____	*U.S. Government* (For Use ONLY by U.S. Government Agencies) [ ] Departmental File [ ] Secret File

Conveying Party(ies)

[x] Mark if additional names of conveying parties attached

Execution Date
Month Day Year

Name (line 1)	Jean-Baptiste Dumas Milne Edwards	03/16/2001
Name (line 2)		

Second Party

Execution Date
Month Day Year

Name (line 1)	Lydie Bougueleret	01/29/2001
Name (line 2)		

Receiving Party

[ ] Mark if additional names of receiving parties attached

Name (line 1)	GENSET, S.A.
Name (line 2)	a Corporation of France
Address (line 1)	24, rue Royale
Address (line 2)	
Address (line 3)	Paris / France / 75008
	City / State/Country / Zip Code

If document to be recorded is an assignment and the receiving party is not domiciled in the United States, an appointment of a domestic representative is attached. *(Designation must be a separate document from Assignment.)*

Domestic Representative Name and Address

Enter for the first Receiving Party only.

Name	John Lucas, Ph.D., J.D.
Address (line 1)	GENSET CORPORATION
Address (line 2)	10665 Sorrento Valley Road
Address (line 3)	San Diego, CA 92121-1609
Address (line 4)	

FOR OFFICE USE ONLY

Mail documents to be recorded with required cover sheet(s) information to:
Commissioner of Patents and Trademarks, Box Assignments , Washington, D.C. 20231

FORM PTO-1619B
Expires 06/30/99
OMB 0651-0027

Page 2

U.S. Department of Commerce
Patent and Trademark Office
PATENT

Correspondent Name and Address

	Area Code and Telephone Number	1-858-597-2600

Name	John Lucas, Ph.D., J.D.
Address (line 1)	GENSET CORPORATION
Address (line 2)	10665 Sorrento Valley Road
Address (line 3)	San Diego, CA 92121-1609
Address (line 4)	

Pages

Enter the total number of pages of the attached conveyance document including any attachments.   #   3

Application Number(s) or Patent Number(s)

☐ Mark if additional numbers attached

Enter either the Patent Application Number or the Patent Number (DO NOT ENTER BOTH numbers for the same property).

Patent Application Number(s) | **Patent Number(s)**

09/663,600

If this document is being filed together with a new Patent Application, enter the date the patent application was signed by the first named executing inventor.

Month   Day   Year

Patent Cooperation Treaty (PCT)

Enter PCT application number only if a U.S. Application Number has not been assigned.

PCT [ ] PCT [ ] PCT [ ]
PCT [ ] PCT [ ] PCT [ ]

Number of Properties

Enter the total number of properties involved.   #   1

Fee Amount

Fee Amount for Properties Listed (37 CFR 3.41): $ 40.00

Method of Payment:
Deposit Account   Enclosed [x]   Deposit Account [ ]

(Enter for payment by deposit account or if additional fees can be charged to the account.)

Deposit Account Number: #   50-1181

Authorization to charge additional fees: Yes [x]   No [ ]

Statement and Signature

To the best of my knowledge and belief, the foregoing information is true and correct and any attached copy is a true copy of the original document. Charges to deposit account are authorized, as indicated herein.

John Lucas, Ph.D.		30/Apr/01
Name of Person Signing	**Signature**	**Date**

RECORDATION FORM COVER SHEET
CONTINUATION
PATENTS ONLY

Conveying Party(ies)

☐ Mark if additional names of conveying parties attached

Enter additional Conveying Parties

		Execution Date
		Month Day Year
Name (line 1)	Aymeric Duclert	01/18/2001
Name (line 2)		
		Execution Date
		Month Day Year
Name (line 1)		
Name (line 2)		
		Execution Date
		Month Day Year
Name (line 1)		
Name (line 2)		

Receiving Party(ies)

☐ Mark if additional names of receiving parties attached

Enter additional Receiving Party(ies)

Name (line 1)		
Name (line 2)		
Address (line 1)		
Address (line 2)		
Address (line 3)		
	City	State/Country Zip Code

☐ If document to be recorded is an assignment and the receiving party is not domiciled in the United States, an appointment of a domestic representative is attached. *(Designation must be a separate document from Assignment.)*

Name (line 1)		
Name (line 2)		
Address (line 1)		
Address (line 2)		
Address (line 3)		
	City	State/Country Zip Code

☐ If document to be recorded is an assignment and the receiving party is not domiciled in the United States, an appointment of a domestic representative is attached. *(Designation must be a separate document from Assignment.)*

Application Number(s) or Patent Number(s)

☐ Mark if additional numbers attached

Enter either the Patent Application Number or the Patent Number (DO NOT ENTER BOTH numbers for the same property).

Patent Application Number(s)			Patent Number(s)		